R E P O R T R E S U M E S

ED 014 734 CG 000 306

THE MMPI AS A MEASURE OF TREATMENT EFFECTS IN VOCATIONAL REHABILITATION. FINAL REPORT.

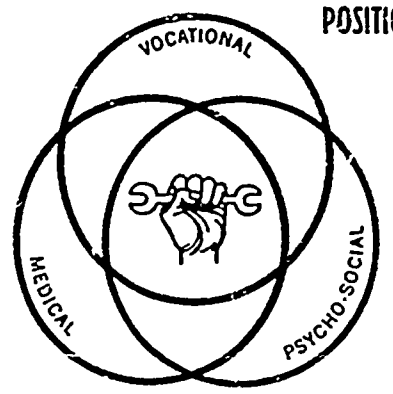
BY- COPELAND, WILLIAM C. AND OTHERS
MINNEAPOLIS REHABILITATION CENTER, MINN.

PUB DATE MAY 66

EDRS PRICE MF-$0.50 HC-$3.32 81P.

DESCRIPTORS- *VOCATIONAL REHABILITATION, *EMPLOYMENT, *PERSONALITY ASSESSMENT, *PREDICTIVE VALIDITY, MINNESOTA MULTIPHASIC PERSONALITY INVEN., LAMBDA MODEL, SUM OF DIFFERENCES, CONFIGURAL SCORING METHOD

THE MINNEAPOLIS REHABILITATION CENTER (MRC) STUDIED MINNESOTA MULTIPHASIC PERSONALITY INVENTORY (MMPI) SCORES TO SEE IF THEY RELIABLY PREDICTED EMPLOYMENT OUTCOME AND TREATMENT EFFECTS. IT WAS DEEMED IMPORTANT FOR AGENCIES TO HAVE A METHOD OF DIAGNOSING PROBABLE RETURN-TO-WORK CLIENTS BEFORE SELECTING THOSE FOR REHABILITATION. A METHOD OF EVALUATING REHABILITATION SERVICES THROUGH MEASUREMENT OF CLIENT CHANGE WAS DEVELOPED. THE MMPI WAS ADMINISTERED TO THREE GROUPS (PILOT, CONTROL, AND CROSS-VALIDATION) UPON REFERRAL TO THE AGENCY AND FIVE WEEKS LATER. STATISTICAL ANALYSIS OF THE TEST RESULTS, MEDICAL AND DEMOGRAPHIC INFORMATION, AND THE WORK SITUATION OF THE CLIENTS PRODUCED NEGATIVE RESULTS. THE PROJECT DIRECTOR BELIEVES THAT THE STUDY FAILED TO PRODUCE RELIABLE RESULTS BECAUSE NONE OF THE PREDICTORS WERE DIRECTLY RELATED TO GOING BACK TO WORK. PROBLEMS OF GROUP SELECTION AND VARIABLE DEFINITION ARE DISCUSSED AS POSSIBLE REASONS FOR THE RESULTS. GUIDELINES FOR FUTURE RESEARCH ON SUCH PROBLEMS ARE SUGGESTED. THERE IS NOT YET AN ORGANIZED SYSTEM FOR MEASURING BEHAVIORS WHICH ARE IMPORTANT IN GETTING AND KEEPING A JOB. (NS)



THE MMPI AS A MEASURE OF TREATMENT EFFECTS IN VOCATIONAL REHABILITATION

Principal Investigators:

WILLIAM C. COPELAND, Biostatistician
American Rehabilitation Foundation

CHARLOTTE F. KAUPPI, Clinical Psychologist
Minneapolis Rehabilitation Center

Project Director:
ROBERT A. WALKER, Assistant Director
Minneapolis Rehabilitation Center

MINNEAPOLIS REHABILITATION CENTER, INC.
1900 Chicago Avenue
Minneapolis, Minn. 55404

FINAL REPORT
May 1966

This investigation was supported, in part, by a reserach and demonstration grant, number RD-1598-P-65, from the Vocational Rehabilitation Administration, Department of Health, Education and Welfare, Washington, D.C.

I. ACKNOWLEDGEMENTS

This one-year project was supported, in part, by a grant (No. 1598) from the Vocational Rehabilitation Administration, Department of Health, Education and Welfare, Washington, D.C., and was a cooperative effort involving the Minneapolis Rehabilitation Center (MRC) and the American Rehabilitation Foundation (ARF). The MRC would like to extend its thanks to the Foundation for providing key services on the project, through Mr. William C. Copeland of its Statistical Services Department. Such assistance was crucial to completion of the project, since the project's author and MRC Research Director, Mr. Richard Wimsatt, left the agency just prior to the project being funded. Apart from Mr. Copeland, thanks should also be given to other American Rehabilitation Foundation staff, such as Iver Iversen, statistician, and Floyd Ayers and Gary Athelstan, psychologists, for their constructive comments on the manuscript; and to David Thompson, who performed numerous computations and tabulations. Final kudos go to Joseph Shechtman for his editorial skills and to the Foundation's financial unit, which authorized services beyond the amount set in the budget.

MRC staff member Charlotte Kauppi contributed the main effort for the MRC and was aided by Jean Orthel, in data collection, and Gary Prazak who contributed his skills in a general review of this manuscript.

Although the final editing was the product of one hand--Mr. Copeland--the actual writing of the chapters was the product of many. Robert A. Walker wrote the summary. The section on background, purposes and procedures was produced by Mrs. Charlotte Kauppi and Mr. William C. Copeland. The technical report was written by Mr. Copeland, with Mrs. Kauppi and Mr. David Thompson; and the section on significance of the results was written by Mr. Copeland.

Our appreciation is also given to the Minnesota Division of Vocation Rehabilitation whose efforts in obtaining a control group and client follow-up data were marked by persistence and skill in the face of many obstacles.

Final thanks go to our clients who spent a significant part of their lives taking what must have seemed to be an endless series of MMPI's.

II. SUMMARY

Purpose:

The purpose of this study was to find out whether the Minnesota Multiphasic Personality Inventory (MMPI) could be used to measure the treatment effects of a vocational rehabilitation center program and whether variables from the MMPI and other services might be used to predict employment outcome.

Procedures:

An earlier study done at the Minneapolis Rehabilitation Center (MRC) showed that changes did occur on MMPI testings. However, because the measure of change was not adequate, and because there was no control group, it was not possible to state that these changes did occur because of the MRC services. To further investigate the problem and do other studies as well, a study of the topic was initiated, which was financed by a grant from the Vocational Rehabilitation Administration. Three separate populations were formed:

> Pilot: This group of 101 DVR referred clients was used in the earlier study. A pre-service and post-service MMPI had already been given. Further work with this group was to involve a third MMPI testing to see if the changes first seen actually held up, the collection of outcome employment data, and the use of demographic and MMPI data to predict employment outcome.
>
> Control: This group consisted of 29 clients who were referred to the MRC and were given two MMPI's without any MRC services between testings. The purpose was to compare MMPI results with the sample which received MRC services to see if MMPI changes occurred spontaneously.
>
> Cross-Validation: This group of 40 clients was used to see if results in the other samples held up under cross-validation.

The first problem encountered was that of obtaining a third MMPI on the Pilot group. After a number of attempts at bringing the group back to the MRC for a third testing failed, this portion of the study was dropped. Securing outcome measures of employment was also difficult. Many clients would not respond, and those who did respond did not report complete information. Finally, an 80 percent sample was obtained - but the questionnaire used had to be shortened. This meant that, instead of having elaborate employment criteria measuring such things as the quality and the quantity of employment, we ended up with a simple measure of whether or not the client was working at the time we got the information.

Attempting to measure MMPI changes was the fundamental problem, and since no one acceptable method is currently in use, three different procedures were used. One consisted of a Sum of Differences score. To get this figure, we subtracted the second testing on each MMPI scale from the first, and added up the figures for every scale, for a given client. The second

system, the Sum-of-Squared-Differences, involved the same procedure except that the differences were squared. The third procedure involved a Configural Scoring Method. This system used the notion of "critical changes", which were defined as those scores which crossed the boundary of a T-score of 70 on the MMPI between the first and second testing.

A final problem was the choice of a statistic to relate the employment outcome data to a set of predictors which included not only the MMPI but demographic and medical data as well. The statistic selected was the lambda model, originated by Guttman. This statistic can be used with several predictors and takes into account the base-rate, or "predicting the mean," phenomenon.

Results:

Several MMPI scales were found to be useful in predicting an employment outcome, when the Sum of Differences method and the Configural Scoring Method were used. However, when these predictors were used in the cross-validation sample they did not hold up.

Further study was then carried out, using other possible predictors such as sex, marital status, medical status, age, etc. Again, although several variables were able to predict working at follow-up, the use of these predictors in the cross-validation group did not hold up.

An earlier study had been carried out at MRC in which the goal was to determine whether or not changes did occur in the MMPI as a result of services. Significant changes in certain scale scores had been taken as evidence for such an hypothesis. However, in a replication study, (not reported here) the changes occurred on an entirely different set of MMPI scales. Our reluctant conclusions were that we could not demonstrate that the MRC services would reliably affect MMPI scores -- and the results of this study provide additional confirmatory evidence for this conclusion.

In summary, we concluded that the results, although promising in the pilot sample, were negative when subjected to cross-validation. Our employment outcome, as used in this unique sample, cannot be reliably predicted by MMPI test results, changes in MMPI scores, and a variety of demographic and other variables.

Comments:

A variety of arguments could be raised by statisticians and experimental design specialists concerning the basic design, sampling procedures, and definition of the measures used. This writer's opinion takes a different direction. My own opinion is that this study failed to produce reliable results because none of the predictors are, in fact, directly related to going back to work. As most Vocational Counselors can testify, people do not fail to get jobs because employers can spot a deviant MMPI during the interview. On the other hand, they usually won't hire someone who whistles in an interview. The MMPI measure of this annoying trait does not exist. Nor will people get jobs if they do not look for work. Again, although high scores on the MMPI depression key might make one suspicious, some pretty unhappy people do

manage to make that first contact before 10 in the morning. It is also obvious to those who do keep jobs that you can believe in the second coming of Christ, hate your mother, and still hold a job. Provided, of course, you don't criticize your boss -- a trait not measured by the MMPI, but still a critical necessity for those who want to keep working.

These comments reflect a basic problem in successful prediction studies in the field of vocational rehabilitation. We do not have an organized system of measuring those behaviors which are important in getting and keeping a job. Instead, we use a psychopathological or medical model, with a few demographic variables thrown in for ease of measurement; then, with a weak set of outcome measures, we hope for good results. A few studies in the field (none, however, which were discussed in this study) have shown such good results. This study shows how such good results might be obtained - just don't do cross-validation.

III. BACKGROUND, PURPOSES, AND PROCEDURES OF THE PROJECT

A. INTRODUCTION

Minneapolis Rehabilitation Center is a private, non-profit, comprehensive out-patient vocational rehabilitation facility. Its purpose is to assess the vocational, social, and emotional problems of the multiply disabled and disadvantaged. Through the use of a simulated work setting (workshop) and the coordinated efforts of an interdisciplinary team consisting of a Vocational Counselor, Work Evaluator, Social Worker, and Clinical Psychologist, and evaluation is made of clients' employability, job readiness, and work skills. Depending upon feasibility, clients are prepared for the labor market and provided job placement services, or they are prepared for and referred to other resources. It is intended that clients will be better able to succeed in employment or to develop a more satisfactory life plan than they have formerly.

Clients range in age from 16 to 65, come from Minnesota and from neighboring states, belong to every socio-economic class, and manifest many types of disabilities and problems. They are seen at a time when they appear to be medically able to consider employment or training, but are vocationally handicapped due to any one of a combination of physical, emotional, intellectual, and educational problems. They are referred by many agencies, including the Division of Vocational Rehabilitation of the State of Minnesota, the OASDI Disability Determination Unit of the State of Minnesota, the Department of Employment Security, and others.

B. HISTORY AND PURPOSE OF THE PROJECT

The Minnesota Multiphasic Personality Inventory (MMPI) is a widely used psychological instrument which was originally developed[1] to discriminate various psychiatric groups from each other and from normals. Since its development, it has come to be used much more broadly. It is used to evaluate personality functioning in the individual case, as well as to discriminate between groups; it is used with normals as well as with psychiatric groups, and it is used in many settings other than hospitals and mental health settings. For example, it is extensively used in educational and vocational settings. The Minneapolis Rehabilitation Center (MRC) is one such setting.

The information from the MMPI is used by the clinical psychologist, in conjunction with results of other tests, interviews, and social history, to relate personality functioning to a client's adjustment in the MRC program, to make predictions about work behavior, to suggest treatment approaches, and to assist in vocational planning. MRC clients usually take the MMPI

[1] Welsh, G.S., and Dahlstrom, W.G., eds., Basic Readings on the MMPI in Psychology and Medicine. Minneapolis, University of Minnesota Press, 1956.

early in the first week of their program. The average length of time they are in the program on a full-time basis is five weeks. When they complete this part of the program, they begin job-seeking or training, wait for the beginning of the training course they have select d, return home, are referred to a sheltered workshop, etc. Early in 1962, it became the practice to administer both "entry" and "exit" MMPIs to clients in order to re-assess personality functioning after MRC services.

In the past, the MMPI has been used at MRC as a clinical instrument. It was being assumed that, in our client population, personality components were being measured by the MMPI and, further, that personality changes would also be reflected on the MMPI (over two or more administrations of the test). Such changes are also believed to be vocationally significant.

Given the assumptions about the relationship between treatment, personality change, and vocational outcome--and the parallel assumptions that personality states, and changes therein, can be measured by the MMPI--the MRC staff decided to compare entry MMPI results with results from exit MMPIs. In a pilot study carried out at MRC in 1963, entry and exit MMPI scale scores (K-corrected T-scale) were compared for 101 clients (the pilot sample). The subjects were selected on the basis of whether they

(a) had both an entry and exit MMPI profile on file;

(b) entered the program between January, 1962 and April, 1963;

(c) could read at about the fifth-grade level or better; and

(d) had been in the program at least four weeks.

This group, which amounted to about 25 percent of all DVR referrals admitted to MRC program services during the time period defined, cannot be considered in any sense "representative" of the whole client group seen during the period (the educational requirements, and the requirements for being able to take a valid MMPI insured that the study groups would be, to some unknown degree, "above average" for the total MRC population.) However, the purpose of the study precluded any possibility of such representativeness, since only a subset of vocational rehabilitation populations are sighted and can take a valid MMPI.

The pilot study established that MMPI scores are sensitive to pre-post program changes (see reference 1). The next question was what relationship did such score change have to actual changes in personality dynamics or behavioral characteristics of the clients over the period of vocational rehabilitation? A way of getting an answer to this question can be outlined here.

In the original application, two things were promised -- a measure of psychological change under rehabilitation and a conclusion as to the effectiveness of vocational rehabilitation in causing psychological change.

We cannot, however, measure changes in someone until we have a measure of change that we trust (in the psychologist's language, a valid and reliable

measure of change). For example, imagine the situation that we had no concept of measured length, and someone came to you and said -- "your child is two inches taller than last year." Unless you had already accepted the measure dimension implied by "two inches taller" (i.e., there is a useful convention called "taller than" and there is an agreed-upon metric which has a stable relationship to the area of reality that you apply it to -- that is, that the relationship between your yardstick and the rectangular opening through which you drive your car every morning and night doesn't change), you would question the odd jargon which he was using. But you don't question it because you have depended upon the measure and your car's fenders are relatively unscathed.

Suppose then that we had a measure of "psychological state" like this, with a measure of psychological change constructed from two readings on this state at different times. Suppose that it "behaves well" with respect to the way other constructs behave, and we trust it in the same way we trust in our measurements of change in length. Suppose, for example, that we know someone whose reading on this measure of psychologic state is "medium" or "high" after vocational rehabilitation, and the following relationships hold true: he has gone back to work and he was out of work when the measure of his psychological state was "low;" his relationships with his wife and children are reported to be "just fine now," and they were not good when his reading was low; and his relationships with his foreman are such that it is reported "the foreman is a bear, but the client seems to be able to get the work done so that the foreman is happy," and, previously, with the low reading, he was reported to be a hothead who had struck superiors with no provocation -- and so on, then we begin to trust the measure of psychological state, and the corresponding measure of change.

Once we had that measure, and trusted it in the same way we trust measurements of length, we could then perform an experiment to test how well MRC is doing. We would take measurements at the beginning on clients entering the program, take measurements on the same clients at the end of the program -- then announce, "this client is being discharged as 'much improved', but some other client had no improvement; indeed, his psychological state measurem is now such that he should be referred to a mental health center for outpatient psychiatric care." At the same time, since we know that people change somewhat in their measurements on this dimension without any special intervention, we also take a control group out of exactly the population we are dealing with in the program, do not intervene with them, then compare the changes in each group. If the changes in the experimental group are such that we can say that the difference between their improvement and the control group's improvement is greater than could be expected by chance, we accept this as some plausible evidence that the MRC treatment causes improvement. If we can depend upon the measure in a precise way, then we can even give the amount by which the clients were improved by the treatment at MRC.

However, we don't yet have that measure of change. So, before we can ascertain how well the measure says MRC is doing, in some scientifically-acceptable evidential fashion, we must first define the measure, relate it to one or more "ultimate criteria" in a variety of ways, and show it to be

useful. Only then can we use it as the basis for a decision on what MRC is doing, in terms of that measure. We can't do both tasks at once, and we must do the one before the other. That is, in the psychologist's language, we must show that a measure is valid and reliable before we can use it to measure something else.

As yet, no one has defined a good measure of psychological change which is based upon the MMPI. Therefore, before we could design and carry out a study which would allow us to make some decision about the effects of MRC treatment upon the psychological state of the client, we had first to develop an adequate measure of that state.

Therefore, in carrying out this study, we "chose" to do what we had to do anyway. We defined a number of measures of psychological state, based upon the MMPI, defined measures of change in the state, and related these changes to a particular "ultimate" criterion.

If the measure had been shown to be related to the criterion defined here, according to the criteria of "adequate prediction" we set out in the study, we could have then tried it again against several other criteria. If it had continued to "work," in a precise and accurate fashion, we would have then declared that we had a measure of psychological change.

At that point, and only then, would we have been in a position to set up an experiment which would have allowed us to measure the effects of MRC intervention upon the psychological state of our clients. That is, we would have had a psychological measure of state for which we know what a reading "meant" (i.e., we would know when a reading was "good" or "bad," greater or lesser, or whatever).

While the approach discussed above provided a rationale for solving the problem of deriving a measure of change from its relationship to a criterion, certain other design problems remained. One particularly difficult problem was that of differentiating between measures of change for which we can discriminate effects of rehabilitation from "random changes" and those for which we cannot do this.

This is the classical experimental problem of differentiating "spontaneous" changes from those caused by some agent or technique of interest. In a laboratory environment, with a prior hypothesis to be tested, such a problem could be solved by use of a simple before-and-after design. That is, define a population of "subjects needing intensive vocational rehabilitation services." Select a sample of such subjects from the population. Allocate them to an "experimental" group which is to receive treatment, and a control group from which treatment is to be withheld. Take the relevant measurements from each group. Then, using a pre-selected criterion, decide whether the treatment group has improved more than the control group.

However, there are important differences which arise when human groups form the population, as they do here. Chief among these is our ability to maintain experimental control. That is, specific to treatments such as vocational rehabilitation (VR)--which is administered in an uncontrolled environment--is that "treatment" might not really be withheld from a "control"

group. For example, suppose that a group were selected for an "experiment," in which the whole group was measured on the MMPI at the beginning and at the end of a five-week period, then measured some time later for their outcome. One-half of the group would be given vocational rehabilitation treatment durirg the five-week period and one-half would not be given such treatment. A number of questions would arise--including some about how we would choose to guarantee that the control group did indeed not receive VR treatment elsewhere during the five-week period, or following the five-week period but prior to the time of outcome measurement. Further, if we attempted to make special social or economic arrangements for the control group, in order to guarantee "no VR treatment elsewhere," how are we to answer impertinent questions about the possible "treatment effects" which are induced by such arrangements?

Along with the question of whether we can withhold VR treatment is that of whether we should. That is, it may be argued that, if we already have an efficacious "treatment," it is unethical to withhold it from those who may benefit. Therefore, the concept of a control group which receives no treatment or a placebo is ethically obnoxious. This problem has been discussed in the literature of drug trials, with the following kind of rebuttal: To insist upon a control group (i.e., a basis for deciding whether treatment group differences are really attributable to the treatment) makes things momentarily difficult for a small group; however, to insist on giving the treatment to everyone precludes the possibility of obtaining confirmation or disconfirmation of the treatment's efficacy, and makes it possible that a nonefficacious treatment will be administered to large numbers of people while, at the same time, progress in research may be stopped (because the "problem has been solved"). Thus, the whole concept of clinical trial is one of selecting the lesser of two unpleasant courses. The analogue of this argument could be given for vocational rehabilitation and, as the possible efficacy of vocational rehabilitation "treatments" becomes more of an issue, such arguments may be expected. In this study, we begged the issue by taking advantage of the fact that--because of an intense shortage of vocational rehabilitation professionals in the State of Minnesota--the interval between DVR referral of clients to MRC and the point at which they could be accepted for services had become about equal to the length of time taken for MRC treatments (five weeks for most clients).

The control group was therefore selected from clients in the MRC waiting line, since an MMPI could be administered at referral from DVR, then five weeks later at acceptance at MRC. Such a group was manifestly a DVR-MRC population, avoided the ethical problems involved in withholding treatment (since they received treatment after the second administration of the MMPI), and minimized the kinds of alternative services which might be given between MMPI administrations.

Such a solution had a compensating disadvantage, however, since it effectively foreclosed any chance of comparing control group outcome measures with those taken on the experimental group. Therefore, the function of the control group was a reduced one--that of furnishing a check on the measure of psychological change defined in the following way: If the scores on the measure defined differed "sufficiently" between the experimental group and

the control group, then they were potentially useful, and should be subjected to further tests. If they did not differ, they were to be ruled out of further consideration, on the ground that a defined measure should satisfy the minimal requirement that it differs sufficiently when taken on a treatment group and on a non-treatment group drawn from the "same population."

There was yet another difficulty. We did not come to this study with well-defined, already-hypothesized relationships. Much of this study was a hypothesis-generating or hypothesis-seeking effort, rather than a hypothesis-testing one. The number of variables and variable combinations searched in order to find relationships of interest was far greater than the number of subjects available. This means that, even had the original set of subjects been obtained on a basis which satisfied the requirements of relevant statistical models, and had a "satisfactory" control group existed, there would still have been no basis for probabilistically-based inference.

This implied a further compromise--the definition of a cross-validation or replication group. Then, for every measure which passed the tests of the comparison of control group with experimental group and for which relationships of interest could be observed in the experimental group, those relationships could be tested within the cross-validation group.

In summary, we can tabulate the kinds of groups to be defined, the purpose for which each was defined, the data to be collected for each, and the kinds of statistical operations to be performed:

Group	Purpose	Type of Data Collected	Operations to be Performed
Experimental (Pilot)	Search for Relationships Between MMPI and Outcome	Demographic Medical MMPI Outcome	(1) Define MMPI-Related Measures of Change (2) Define Measure of Outcome (3) Find Statistical Relationships between Measures of Change and Measures of Outcome
Control	Find MMPI-Related Measures of Change Which Differ In Behavior in Treated and Untreated Groups	Demographic Medical MMPI	Make Comparisons Between MMPI Measures of Change for Experimentals and Controls for Discriminative Power
Replication	Decide Whether Relationships Found in Experimental Group are "Chance" or "Real" Ones	Demographic Medical MMPI Outcome	(1) Derive Decision Rules For Prediction From Statistical Relationships Found With Experimental Group

C. PROCEDURES

The original pilot experimental group was selected according to the definitions given in the previous section. As noted before, the 101 files selected for this group amounted to about one-fourth of those clients served during the sixteen-month period of Jan. 1962 to April 1963.

As the project got under way, a cross-validation group was constructed by taking all clients served by MRC between April, 1963 and January, 1965, and finding the subset of these which answered to the specifications under which the pilot group was constructed. There were 100 files of sighted clients for which the criteria for inclusion in the study were met. Forty of these were selected randomly, with stratification on sex, the result being two groups for whom identical data were to be collected. The size of these groups, by sex, was as follows:

Table 1

Number of Subjects in Pilot and Cross-Validation Groups by Sex

Group	Males	Females	Total
Pilot (Experimental)	57	44	101
Cross-Validation (Replication)	22	18	40
Total	79	62	141

By definition, data were available on two occasions of MMPI testing. Other file data taken included age at time of entry into MRC, sex, IQ, marital status, education, time in program, and medical status at entry into MRC. These data are tabulated for all groups in the appendices.

Plans had been made to secure a third MMPI score, so that the stability of change scores could be assessed. That is, for those measures of change which turned out to be "adequate," under the tests discussed mentioned in Section IV.B and discussed more fully in Section IV.C, the question of the stability of such changes in the post-rehabilitation environment was considered to be of interest. Given that rehabilitation causes some change in the personality structure, the question arises whether such changes are enduring or transitory. A third MMPI score might have furnished some evidence which would speak to this question.

Further, in order to be able to reconstruct the post-rehabilitation history of the client and to construct a measure of outcome, a 48-item questionnaire was developed, regarding job outcomes, economic outcomes, job satisfaction, social satisfaction, changes in demographic status, and the like (see Appendix V).

In order to collect these data, we intended to bring the pilot group back to MRC for a day of tests and interviews. This plan turned out to be quite unrealistic. By the time the study got underway, some of the pilot subjects had been out of MRC more than three years (the median time since leaving MRC was 31-32 months), and were not interested in returning to the Center to spend several hours in testing and completing questionnaires without some form of compensation for the use of their time. (Three mailings to the 101 clients resulted in only nine clients returning to MRC for tests.) This feeling was even stronger among the fraction (about one-half) of the pilot group which was living outside of the Twin Cities.

Therefore, the plan to secure a third MMPI score was dropped completely, and a mail survey technique was attempted. It quickly became clear that even a 48-item questionnaire was too much for this group. Therefore, the questionnaire was shortened to what were considered the 12 crucial items (see Appendix V). An initial mailing of the shortened questionnaire resulted in a 40 percent response rate. Subsequent mailings, at four-week intervals, to non-respondents brought total response, by all methods, up to about 60 percent.

By this time, it seemed evident that a case-by-case strategy was needed. Other agencies, relatives, friends, or any others who might have knowledge of the client and who might be of help in securing a response to the questionnaire, were contacted. Personal telephone calls were made to all subjects who could be reached by telephone. After reaching a dead end with respect to leads, after being turned down by the client or a relative, or after three contacts without success, the follow-up was discontinued. Ultimately an overall response rate of 80 percent was reached, as indicated in Table 2, below.

Table 2

Response Rate to Outcome Questionnaire, by Group

	Experimental Group					
	Pilot Group		Cross-Validation Group		Total	
	N	%	N	%	N	%
Respondents	84	83.2	29	72.5	113	80.1
Non-Respondents	17	16.8	11	27.5	28	19.9
Total	101	100.0	40	100.0	141	100.0

Of those classed as nonrespondents, it was learned on contact with the subject or a relative that six were employed (two of whom were in the Armed Forces), 5 were unemployed (including two who had been rehospitalized and one who was in a rest home), and one had died. Of the remaining 16, eleven had completely dropped from sight--all letters sent to all addresses tried were returned--and five simply did not respond--their letters were not returned, but they could not be reached by telephone.

Among those classed as respondents, response was not complete for every question on the outcome questionnaire. The implications of such nonresponse are discussed more fully in the section of predictor-criterion relationships in Section C of the "Technical Report. The response rates for individual questions ranged from 50.4 percent of the responders (40.4 percent of all those in both experimental groups) on the question about the amount of hourly pay currently being received, to 97.3 percent of all responders (78.0 percent of the total sample) who replied to the question: "Are you now working?" Results for the entire questionnaire are tabulated below:

Table 3

Number and Percent of Respondents to Individual Items in Follow-up, by Sample Group

Question	Pilot Group (Total N = 101)		Cross-Validation Group (Total N = 40)	
	N	Percent	N	Percent
Are you Now Working?	82	81.2	28	70.0
Number of Hours Worked per Week	46	45.5	19	47.5
Hourly Pay	40	39.6	17	42.5
No. Jobs Held Since Leaving MRC	74	73.3	28	70.0
No. of Months Worked	48	47.5	21	52.5
No. Moves Since Leaving MRC	80	79.2	27	67.5
Change in Marital Status Since Leaving MRC	79	78.2	28	70.0
Time "Laid Up" Since Leaving MRC	76	75.2	28	70.0
Change in Overall Living Status Since Leaving MRC	80	79.2	26	65.0

The problems which led to the particular kind of control group used in this project were discussed earlier. In essence, the plan to secure members of such a group was as follows: Counselors of the Minnesota State Division of Vocational Rehabilitation (DVR) were to administer an MMPI to clients who were prospective referrals to MRC, with a goal of 75 such referrals. Given the expected attrition between referral and entry, a loss of 10-15 subjects could be expected. Given the DVR referral rate to MRC, it could be expected that such a control group would be developed within four to five months.

Three months and three exhortatory letters later, MRC had on file 11 MMPIs from DVR referrals, although the referral rate to MRC had not fallen below the expected rate. Of the 11, six had actually taken an entry MMPI, thus qualifying as control group members. Personal contacts with DVR counselors revealed that it was impossible in many cases, and difficult in most, to do the necessary client testing. Many of the counselors work in small offices with limited testing facilities. Many clients never come into the office; contacts by the counselor are on field trips, frequently to the clients' homes. In such cases the counselor would be reluctant to administer the MMPI and wait the hour or more necessary for the client to finish it. Thus, for a number of practical reasons, it was unrealistic to expect individual DVR counselors to be able to deliver MMPI pre-tests at the expected rate. Therefore, the number of cases in the control group was smaller than expected (29). Details on the demographic, psychological, and medical make-up of this group, which were collected at entry, as with the other two groups, are given in Appendix I.

As a result of this work, the number in each of the three groups, for each of the two stages of information-collecting, was the following:

Table 4

Size of Each Group in the Study--by Group and Stage

Stage of Information Collection	Pilot	Cross-Validation	Control
At Entry Into MRC (Medical, MMPI and Demographic Data)	101	40	29
After Rehabilitation Treatment at MRC (Follow-up Data)	82	28	-

IV. TECHNICAL REPORT

A. INTRODUCTION

One of the original aims of this project was to find an adequate measure of the client's change under treatment, using the MMPI. Much of the problem of candidates for vocational rehabilitation is psychological, and a major component of intensive rehabilitation consists of services aimed at the solution of problems in this area. If the services cause changes in emotional dynamics during the rehabilitation process, the MMPI may provide some measure of these changes.

If such a measure is to be provided, two problems must be solved:

I) Some stable, MMPI-related measure of psychological change must be found.

II) The measure must be established as "relevant" or valid by its relationship to some criterion.

B. DEFINING A MEASURE OF CHANGE

The solution to the first problem can be undertaken in what is essentially an arbitrary way. At this time there is no univerally accepted measure of MMPI-related psychological change. Therefore we are free to choose from any possible MMPI-based measures that may seem useful. This entails choosing from an infinite number of possible scoring systems.*

Since we can only deal with a limited number of such possible measures of change, we must choose those which seem a priori to have a high probability of success in providing a stable, adequate measure of psychological change. If we do not find a satisfactory measure of change, this does not demonstrate that one does not exist--we may not have looked far enough.

1) A simple sum-of-scale-differences type of change measure.

---

*For example, if we choose from a very restricted set of possibilities, defined as the "difference" between two administrations of a test with n questions, assuming only one system of weighting combinations of the items into scales and a single method of differencing between the two administrations, there will be 2n possible score types. With the MMPI, these items are commonly lumped, with some questions showing up in more than one of the scales, into 13 scales (14 for this project). Therefore, if we take only T-scale scores for the two administrations of the MMPI, we have $2^{14} = 16{,}384$ possible ways of getting a difference score.

2) A sum-of-squares-of-scale-differences type of change measure.

3) A configural-scoring method.

These change measures are defined in the following way:

1) The Sum of Differences Method.

Here, we take the T-scale score on the second administration of the test, then subtract it from the score on the first administration. For a given subject, these scores are summed over all 14 scales used in the project (ES, E, F, K, 1-0; the sign of the difference is reversed for ES), and the sum of these 14 differences is taken as the measure of change i.e., if yi is the T-score on the second administration of the MMPI, for the ith scale, and xi is the T-score on the first administration, then the difference score for the jth subject is defined as $d_j = \sum_{i=1}^{14} (x_i - y_i)$. From the measurement point of view, we would want those who had benefited from treatment to have a score, dj, which was large and positive, and those who had no treatment (or those who had had treatment, but were not helped) to have scores which were near zero, or were negative.

2) The d^2, or Sum-of-Squared-Differences, Method.

Rao's generalized distance function, which has been put forward at various times as a way of applying the least-squares criterion to discriminant problems, has some intuitive appeal as measure of change on an MMPI profile score. Such a statistic is a measure of the "volatility" of the change in scores over two administrations in comparison to the sum-of-differences method. This becomes clear when we look at what happens to the changes on each of two scales. If one is large and positive, and the other is large and negative, the sum of the two will be near zero for the sum-of-differences method, but very large for the sum-of-squared-differences method. This method is defined as:

$$d_j^2 = \sum_{i=1}^{14} (x_i - y_i)^2$$

3) The Configural Scoring Method.

As a last way of measuring differences between two administrations of the MMPI, we have investigated configural score relationships. This kind of scoring is motivated by classical MMPI clinical typing, in which various configurations of elevated and non-elevated scores are delineated, which serve as discriminators among various kinds of personality types.

In order to define some useful measurement of change over two administrations, it was felt that one plausible, and not exceedingly complex, way of constructing a measure of change was to use the notion of the critical boundary which arises in many discussions of MMPI scoring. Often, clinical interpreters focus their interest on whether a specific T-scale score is 70 or over (i.e., beyond the two-sigma point on the standardized scales), or not.

On the basis of this notion, "critical changes" were defined as those scores which crossed the T=70 boundary between the first and second administrations. Those which went from under 70 to 70 or over between first and second administrations would be coded as one kind of extreme score, and those which went from 70 or over to under 70 would be considered the opposite kind of extreme score. Scores which did not cross the boundary between administrations would be interpreted as non-extreme. The scores were then coded in the following way, one for each of the 14 scales:

Configural Coding

	Second Administration	
	Under 70	70 or Over
First Administration Under 70	2	0
First Administration 70 or Over	3	1

The scores can be given ordinal meaning, with an "0" signifying the "worst" kind of change, and a "3" the "best" kind of change. Unlike the first two kinds of scoring systems, in which one summed "global" score was used, the measure of change in this scoring system is a set of scores; for example, a subject's change measure might be the following:

ES	L	F	K	1	2	3	4	5	6	7	8	9	0
0,	0,	1,	3,	2,	3,	C,	0,	2,	1,	1,	1,	2,	3)

Such a set of scores would be interpreted as follows:

1) The subject had a score of less than T=70 on the first administration, and a score of 70 or more on the second, on scales ES, L, 3, and 4.

2) The subject had a score of 70 or more, on both administrations, on scales F, 6, 7, and 8.

3) The subject had a score of less than 70, on both administrations, on scales 1, 5, and 9.

4) The subject had a score of 70 or more on the first administration, and a score of less than 70 on the second administration, on scales K, 2, and 0.

C. THE RELATIONSHIP OF THE CHANGE MEASURE TO A CRITERION

The second problem, establishing the relationship as relevant or valid by relating it to some criterion, can be solved by two kinds of operations in

the case of global scores, and one kind of operation in the case of vectors of scores, within the constraints of the sample size available for this study.

1) Checking for "Spontaneous Change"

The first kind of operation is used as a simple check-criterion. In the procedures section, it was noted that a common experience in the analysis of MMPI test-retest data was that of finding a change in taking the test (other than the common factor that the group was taking the MMPI for a second time). Profiles generated from the second administration of the test commonly seemed to indicate an "improved" profile. Therefore, we might require that any MMPI-based measure of change during rehabilitation services should be sensitive enough so that spontaneous change attributable to a second administration of the test would not be large enough to cloak or wash out the measure of real psychological change.

If we suppose, as a preliminary assumption, that the rehabilitation process does indeed "cause" a psychological change, we may then require that the value of whatever measure we have defined have some discernible difference (when it is estimated for a "treated" or experimental group) from the value it takes when another group, equivalent in every respect other than that of having the "treatment", is subjected to the same measurement (the "control group"). If there were no difference between experimentals and controls on the same measure, then the measure would be disqualified for further testing for its possible utility - on the ground that, whatever other uses the measure might have, it does not have the critical property of being able to discriminate treated from untreated.

Accordingly, a "control group" was defined, which was generated by the same referral process as the "experimental group". The time interval between testing was approximately equal (an average time of five weeks) for both groups. The control group differed, however, in having received no services between the time of the two tests, while the experimental group did receive such services. The administrative constraints on the design did not allow for a randomization process in allocating to the two groups, so that problems of generalizability exist. For example, the advantages of recourse to a probabilistic model are lacking. Therefore, whether the experimental groups and the control groups are "equivalent" with respect to the characteristics which might be important in their differential response to treatment is not decidable in any useful sense. Nevertheless, comparison on characteristics considered routinely important in vocational rehabilitation is of interest. Some of these comparisons are discussed as background to the comparison of the groups on the MMPI-based measures of change which we have defined. Also, group means and standard deviations for each of the 14 scales used in this study for each sex, for each of the two occasions of testing, and for each of the three groups studied, are given. These comparisons are included as Appendix I.

The basic method for comparing the two groups is the following. Since we do not know how "good" any score is, we will compare the distribution of scores. That is, we will look at the proportion of those within a group having a

score less than some value k, over the entire range of k, i.e., the cumulative distribution function of the group. We will then compare the distribution functions of the three groups, using a nonparametric procedure for comparing such functions, the Kolmogorov-Smirnov (K-S) statistic. This statistic (the greatest vertical discrepancy between pairs of empirical distribution functions) is based on probability considerations, namely, that we assume a null hypothesis which postulates a common distribution function for the two groups, under the random sampling model. If the two sample groups differ so much that we would believe, on long-run sampling considerations, that this would be a rare event given that the two groups are "really" samples from a common distribution, we then decide that the two groups probably come out of "differing" improvement score distributions. It should be noted that these groups were not derived in such a way that a probability model can be easily assumed; nevertheless, we can use it in an "as if" spirit, as a basis for future work.

Several different versions of the sum and sum-of-squares measures were used, and the analyses were done separately for each sex, as well as for both sexes at once, on the chance that reactions on the measure might be a function of sex as well as treatment/non-treatment. Such an analysis for the configural scoring method was impossible, since experimental and control groups would be needed which would be some multiple of $(4)^{14}$ in size, in order to test whether the two groups differed in distribution or not.

The results, using the K-S procedure to test for differences between experimental and control groups, were essentially negative. Tests were performed for the sum of difference scores and for the sum of squared difference scores. It was our intention to test also for possible sex interactions; however, the number of females in the control group was far too small to allow any kind of useful comparison when a fairly insensitive nonparametric procedure was being used. Figures 1 and 2, which are graphs of the two measures, make clear the fact that the empirical distribution functions of the three groups differ very little, and could have easily come from the cumulative distribution functions of a single underlying population.

Tables of the distribution function of each of these difference scores are included in Appendix II. Although the distribution function of the two major groups does not differ much for the sum of difference scores, Table 1, Distribution of Sum-of-Differences Scores, exhibits some interesting differences between groups when we subdivide the experimental (pilot plus replication) group according to a rough measure of relative work success after rehabilitation. We will discuss the implications of these differences in the next section.

2) Finding Predictor-Criterion Relationships

For those measures of change which passed the first test, a second test was applied.. Was the measure of change related to some "ultimate criterion" or a convenient surrogate therefor?

a) Earlier, it was noted that there is at this time no widely-accepted measure of MMPI-based or MMPI-related psychologi-

cal change. One important reason for such a state of affairs is the fact that there is no "ultimate criterion" of improvement due to psychological change that is widely accepted.

That is, we need some kind of construct to serve as a criterion for the measures of psychological change which will tell us when these measures of change behave in such a way that we can decide that the change in the underlying thing measured (some aspect of the client's personality) is "good" or "preferred" - as against when the change should be taken as an indicator of something "bad" or "not preferred" or "no change."

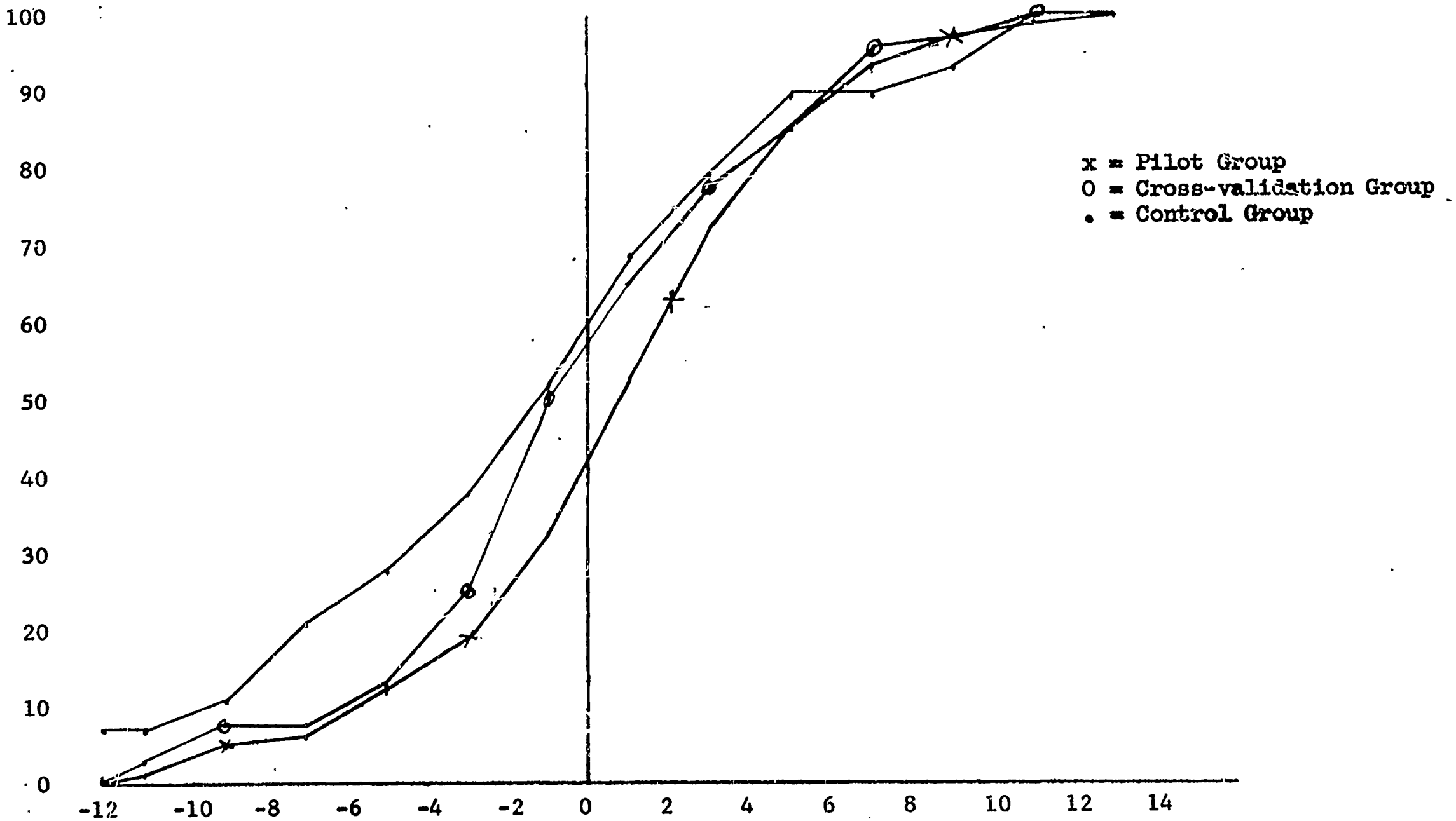
Otherwise we might make all the measurements we please, using any kinds of "measuring sticks" we care to define; but such measurements would not, in themselves, give us any information about the meaning of the changes in the measurements. They would make sense only when they (and the construct of which they are a part) were related in some "lawful," regular way to some other set or sets of measurements (and the construct or constructs of which these other measurements are a part) which already have meaning for us.

If we find a strong relationship between the measure of change and the measurements on the criterion, we can then attribute, in a "backward inference," from the "preferred" states on the criterion back to their related states on the predictor. For example, if we find that virtually all of those whose score on some measure of improvement is equal to or greater than 150 were also in the preferred state on the criterion (for example, went back to work), while virtually all of those with a score of less than 150 on a predictor were in the non-preferred state (didn't go back to work), we would impute to the characteristic of "having a score equal to, or greater than, 150" the associated characteristic of "improvement under treatment."

In psychology, however, there are some special complications. First of all, aside from the bounds given by gross personality disorganization or sociopathic characteristic disorders, what constitutes a "good" behavioral trait and what constitutes a "bad" one is a matter of no little disagreement, resting on valuational grounds which differ much between psychologists, as well as between most other classes of humans. Therefore, what constitutes a change for the "better" or for the "worse" in any absolute sense is bound to be a matter of disagreement.

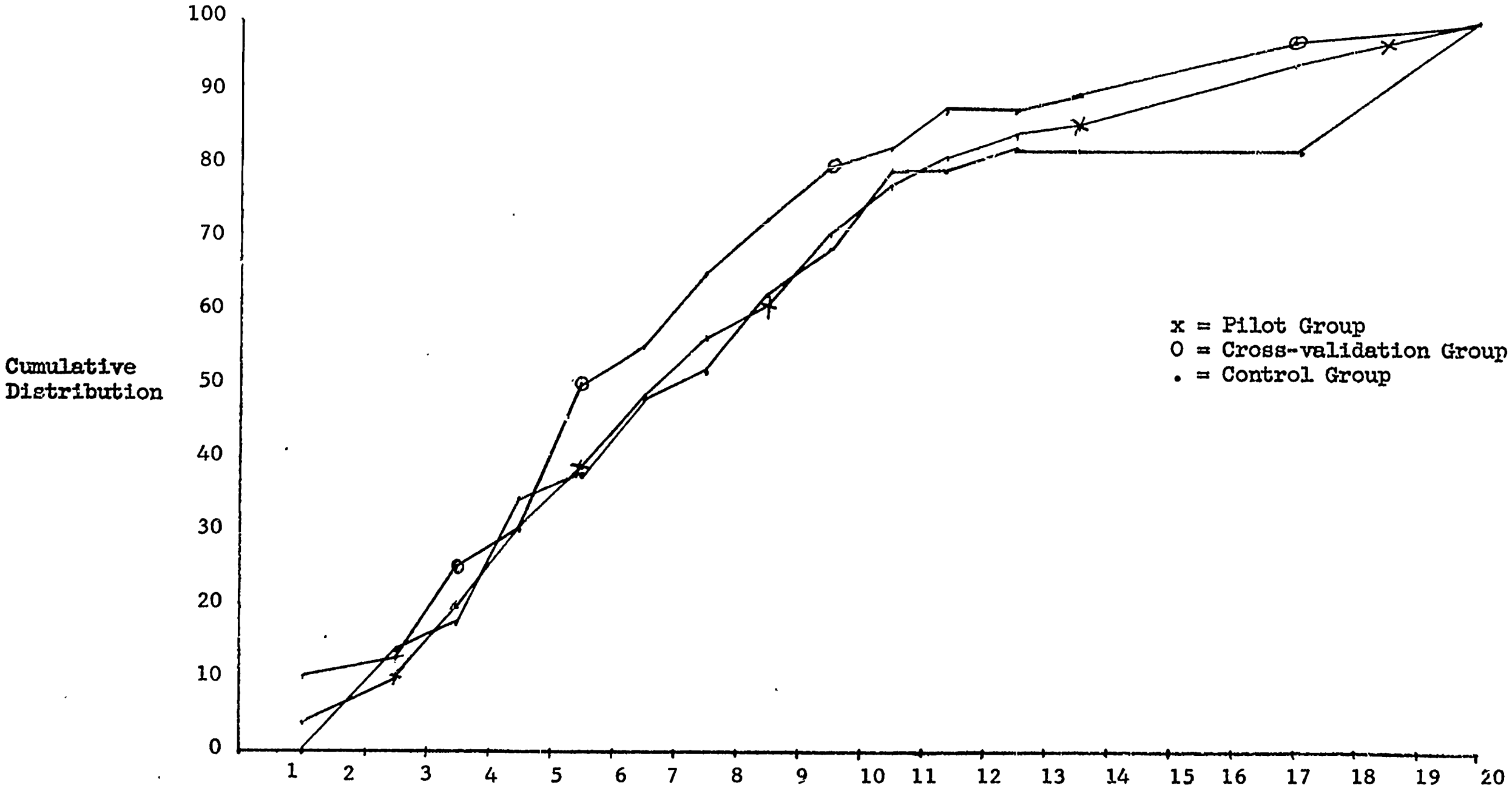
We can, however, usually secure somewhat more agreement by limiting our discussions of "better" or "worse" to sharply circumscribed areas, where more people may be able to agree within the limited context of interest, on what "better" and "worse" might be. This is most easily done within the areas of life in which constructs which specify measures of preferredness exist which are "natural" to the area (e.g. in certain areas

Figure 1: Cumulative Distribution of Sum of Difference Scores on Two Administrations of the MMPI, For Three Vocational Rehabilitation Groups



*Note: Score expressed in tenths

Figure 2: Cumulative Distribution of Sums of Squares of Difference Scores on Two Administrations of the MMPI, For Three Vocational Rehabilitation Groups



*Note: Score expressed in hundredths

of economic life, where the object is to maximize sales or profit or "prestige of the institutions", within certain bounds given by convention, taste, or principle.)

In the area of vocational rehabilitation, we have available a limited but "natural" criterion of preferredness, to which we can (tentatively, at least) relate any defined measure of psychological change. Such a choice can be rationalized on the basis that, while the "therapeutic" aspects of the work of a vocational rehabilitation center are extremely important, they are, in the context of this study, instrumental to an underlying economic purpose. Therefore, the specific criterion of interest in this study must be in some fashion related to the economic outcome of the vocational rehabilitation process.

Such an economic outcome can be defined in a variety of ways, ranging from the simplest kind--whether the client has been accepted on any kind of job for some interval of time--to the most complex--e.g., a measurement of the "quality" of the job or its "satisfactoriness," over some longer interval, using monetary measures of changes in the client's income stream (including both earnings and transfer payments) and measures of outcome for the client's status on these measures prior to rendering the services.

It might be argued that any economically-based criterion is inappropriate to the validation of a measure of psychological change, and that such an objection would apply a fortiori to such a specifically-defined economic criterion. To this, we can only remark that any "ultimate" criterion which would operate over the whole set of human purposes is not definable at the present time (and probably never will be), and that any useful criterion we might define is conditioned by a great number of historical, social, and economic factors. Therefore, any such criterion will be useful only for certain limited purposes (here, for example, for the purpose of finding a way of distinguishing changes for the "better" in psychological state from those which are not "better").

It should be noted that the (limited) possibility always exists that some measure which we find to be related to one kind of a criterion might turn out, on further investigation, to be related to other, entirely different kinds of criteria in sensible ways and thus furnish us with a more or less "universal" measure of say, psychological change. Until we reach such a measurement utopia, however, we shall have to be content with the fact that our measure of change is useful, at most - in the present instance - only "with respect to" the criterion to which it has been related.

Because of the administrative and budget problems involved in

collecting adequate kinds of information from the experimental groups, the actual criterion measurement finally used in this study was on the low end of the complexity spectrum. Originally, it was planned to secure data concerning the amount of work done since leaving the rehabilitation program, and to reconstruct total earnings, so that quantitative measures of "amount" and "quality" of work following rehabilitation would be available as a measure of "total improvement."

However, because of the low budget for case-finding and follow-up, relative to the amount of time and effort needed to find the experimental group at follow-up, the kind of near-continuous follow-up monitoring needed to construct such estimates could not be undertaken. Because of problems of mental retardation, neurological problems affecting remote memory, and the sheer inability of many clients to reconstruct a description of what was sometimes a patchwork of jobs since leaving rehabilitation services, estimates of amounts of time worked since leaving the program, as well as the amount of money earned, were considered too unreliable to furnish useful data upon which estimates of total improvement could be based (specific data are included in Section III. C.). As a compromise, the simplest dependable measures available were taken: Had the client ever worked since treatment? Was he working at the time of follow-up?

Of the two measures generated by these questions, the one concerned with work status at time of follow-up was considered better than that associated with whether the client had ever worked since his receipt of rehabilitation services. Virtually all of those who had received MRC services had worked at one time or another (85 percent) since leaving MRC, thus making this measure less useful for discriminative purposes. On the other hand, 54 percent of the experimental group had been found working at the time of follow-up a much more useful percentage for discriminative purposes. However, discriminative purposes are not the prime criterion here--the real question is whether those found working at follow-up can be considered a separate, well-defined group, and whether this group can be considered "different" from those not working at the time of follow-up.

Such a question can be answered in favor of a measure of those working at follow-up in this way: Few of the clients (roughly about 10 percent) appear to have worked more than 90 percent of the time since they left MRC. Thus, at times, even some of the "best" graduates of the program are currently out of work. Therefore, at any one point in time that a follow-up survey is carried out, there is a non-zero probability that some of these "best" clients will show up in the "currently not working" group. At the same time, some of those graduates who have worked hardly at all, but have worked a little, may be working at the time of follow-up and may be misclassified as "more improved", since they show up in the group labelled "currently

working".* However, the probabilities are much higher that those who have the best work records will show up in the "not currently working" group. Therefore, in this study, we have used the "currently working" versus "currently not working" dichotomy in a number of crucial areas.

An analysis of the data having to do with the percentage of time worked by each client since leaving MRC confirms this point of view. Of the 81 persons for whom there were data in experimental group I, 42 were working at follow-up, 26 others had worked at some time since leaving MRC, but were not working at follow-up, and 14 clients had never worked since leaving MRC,(see Table 5). Data were not available on amount of time worked for 14 of the 42 working at follow-up and for 7 of the 26 who had worked since leaving MRC, but were not working at follow-up; therefore, percentages of time worked had to be imputed for those for whom no data existed. Depending upon the kind of estimator used (median or unweighted mean of percentages) and the kind of imputation made for missing data (worst assumption: "those with missing data worked zero percent of the time"; best assumption: "those with missing data worked 100 percent of the time"), those found to be working at follow-up had worked, "on the average," 43 to 90 percent of the time, with a "best estimate" of 55-67 percent of the time, since leaving MRC. Using the same approach for those who were not working at follow-up, but had worked since leaving MRC, we find that these had worked, "on the average," from 22.5 to 46.5 percent of the time, with a "best estimate" of 27.5 to 29 percent of the time, since leaving the rehabilitation program.

Since every estimate for the percentage of time worked since leaving MRC for the "now working" group is superior to every estimate for the "worked but not now working" group, except for the case in which we make the very best of all possible assumptions for these not now working and, simultaneously, the worst of all possible assumptions for those currently working, we conclude that errors of misclassification which might arise from using the "now working" group as our "superior," criterion group will be moderate.

If we consider only the groups for whom data on percent of time worked since leaving MRC are available we can make an estimate (perhaps heavily biased) of the size of the possible misclassification error due to using the "now working" criterion. This error would appear to involve no more than 25 percent of the group for whom data are available (i.e., less

*For a concise outline of the problem of probabilities of misclassification and their costs, and some useful solutions, see Cronbach, L.J., and Gleser, Goldine G., Psychological Tests and Personnel Decisions, pp. 1-29, 2nd Edition, University of Illinois Press, 1965.

than one-fourth of the "now working" group had worked less than 40 percent of the time since leaving MRC, and less than one-fourth of the "not working, but have worked" had worked 40 percent of the time or more since leaving MRC; these groups constituted only one-seventh of the total "now working" group--after the "no percent-of-time-worked" data were included--and one seventh of the "ever worked" group).

While the problems of misclassification occasioned here are not crucial, the problem should nevertheless be taken into consideration in any future planning of studies of this type. A consideration of the probabilities of being found at work at any given moment, given that one is a non-successful-or successful-graduate of a vocational rehabilitation "treatment" argues for monitoring post-rehabilitation employment at least twice in any kind of follow-up which is intented to yield serious conclusions. In this way, the probabilities of misclassification of clients with respect to their post-rehabilitation success will be minimized, even if information is secured only at the point of follow-up (of course, the follow-up time points need to be judiciously selected).

Table 5

Percent of Time Worked Since Leaving MRC for Pilot Group Follow-up Respondents, By Work Status at Time of Follow-up

% of Time Worked	Now Working	Not Working, Have Worked	Worked	Total
No Information	14	7	-	21
0.0	-	-	14	14
0.1 - 9.9	2	5	-	7
10.0-19.9	2	1	-	3
20.0-29.9	1	4	-	5
30.0-39.9	1	5	-	6
40.0-49.9	3	1	-	4
50.0-59.9	6	1	-	7
60.0-69.9	1	1	-	2
70.0-79.9	2	1	-	3
80.0-89.9	3	-	-	3
90.0-99.9	7	-	-	7
Total Respondents	42	26	14	82

b) Choosing a Measure of Relationship

When we choose a measure of relationship for the conditions involved in this problem, there are some difficulties involved. First of all, we are dealing (possibly) with more than one variable predicting to a criterion; the criterion is a non-quantitative dichotomy (working at follow-up versus not working at follow-up); and, finally, the predictor variables are polytomous (either dichotomous or quartotomous in the problems dealt with here). This implies that we will be dealing with a contingency table in n dimensions (with n the number of predictor variables utilized). Prior to the mid-1950's, no theory was available to allow us to deal with an n-dimensional contingency table in such a way that we could make some measure of the degree of association. And, until the last five years, no theory and associated methods were available such that sampling problems could be dealt with (the early work was concerned only with the case in which we are dealing with the whole population). Today, there are several methods available. Of these, a useful model seems to be the lambda model, originated by Guttman and developed by Goodman and Kruskal. Essentially, this is a predictive model, which is strongly analogous to a multivariate linear regression model.

We can define the model in the following way. If we have N subjects for whom we are making a prediction of whether they are A or B, we have the following kind of matrix:

		PREDICTION		
		A	B	
ACTUAL	A	n_{11}	n_{12}	$n_{11}+n_{12}$
	B	n_{21}	n_{22}	$n_{21}+n_{22}$
		$n_{11}+n_{21}$	$n_{12}+n_{22}$	$N = n_{11}+n_{12}+n_{21}+n_{22}$

When we make the prediction, we will predict that some people are A and some are B. Some of these predictions will be correct ($n_{11} + N_{22}$ out of N of them; i.e., n_{11} were predicted to A's and were A, n_{22} were predicted to be B's and were B); some will be incorrect ($n_{12}+n_{21}$ out of N of them; i.e., n_{12} of them were predicted to be B's, but were A, and n_{21} of them were predicted to be A, but were B).

Now the proportion of correct predictions ($n_{11} + n_{22}/N$) is what we will define to be our HIT RATE.

As we look at this measure, two things are obvious: 1) If we

flip a fair coin as each person comes up for prediction, we will predict about 50 percent correctly (i.e., $(n_{11}+n_{22})/N = .50$) in the long run. This level of prediction we call CHANCE (50:50) PREDICTION. Therefore, any information which would lead us to predict worse than 50 percent correctly in the long run is worse than no information at all; we would be better off flipping a coin. This means that any information we use would be useful only if it helped us to do better than chance prediction. 2) There is one piece of information which is especially helpful in allowing us to predict correctly. Suppose that we know before predicting that 65 percent of the people are always A. Then we could simply predict, for every person for whom we must predict, that he will indeed be A. We will then be correct 65 percent of the time. The level of prediction attainable by knowing this one piece of information, the level of average occurence, we call BASE-RATE PREDICTION. It is sometimes called predicting the mean. It should be easy to see that if, in actuality, A's and B's are divided 50:50; then chance prediction and base-rate prediction will give the same results; but, as the actual proportion of A's departs from the 50-percent level, the hit rate results obtainable from base-rate prediction become better and better.

Therefore, while hit-rate is a good measure of success, it is not entirely satisfactory, since we can do no worse than 50 percent correct at any time, and no worse than the base-rate, if the base-rate differs from 50 percent. A more satisfactory measure of our success in prediction would be a measure of how much better we do than the level achievable by chance or the level achievable by predicting the base rate.

Such a measure is LAMBDA B. If we have information on which we are basing our prediction, we may measure the value of our prediction by use of Lambda B, which is defined as: "The decrease in the probability of predicting incorrectly, as we add information". Then, if a new piece of information (over and above, say, the base-rate) does not allow us to predict more accurately than if we just knew the base rate, then it is not useful information.

All of the measures defined here are given their definition for population (as opposed to sample) statistics in an article by Goodman and Kruskal in the Journal of the American Statistical Association, pp. 732-764, December 1954. Asymptotic distribution theory is derived for these statistics in a 1964 article by the authors in the same journal.

Under the conditions of this study, the relationship of base rate, hit rate, and lambda is the following:

Table 6

Relationship Among Three Predictive Measures, MRC-MMPI Study, for Pilot Experimental Group (N=82)

Number Guessed Correctly	Base Rate	Hit Rate	Lambda
42	.512	.512	.000
45	.512	.549	.075
49	.512	.598	.175
53	.512	.646	.275
57	.512	.695	.375
61	.512	.744	.475
65	.512	.792	.575
69	.512	.841	.675
82	.512	1.000	1.000

As is clear from Table 6, lambda takes values from zero to 1.000, being equal to zero when additional information allows us to predict no better than the base rate, and 1.000, when additional information allows us perfect prediction, moving between zero and one at a uniform rate between the base-rate number and the total-group number. This procedure has other desirable properties, which are discussed in the basic article in which it is defined.

c) Statistical Results

The check on the relationship between controls and the two experimental groups showed that the sum of squared differences score seemed to be of little use in discriminating the two groups. The sum-of-differences score, while of little use for total groups, could be shown to be discriminative in other ways. In this section, we discuss the results of relating this latter score and the configural scoring methods to the job-outcome criterion.

(1) Sum of Difference Scores Method. Although total experimental and control group differences were of no interest, the distribution of scores by the work status of the experimental group versus the distribution of the control group scores was of interest. The data in Table 7 below would seem to imply a relationship between score and quality of treatment outcome for those in the experimental (pilot plus replication) group. That is, the worse the post-rehabilitation work record (or, if no data could be

secured, and we assume the post-rehabilitation work record of the "no data" group to be much worse than the average--so that we assume that the no-data group is more like the "no work" than the "work" group), the greater the proportion of "bad" (negative) difference scores (over two administrations of the MMPI). Further, the score distribution of the "untreated" (control) group seems to be more like that of those who were "unsuccessfully treated" (those with bad work records) than those who were "successfully treated" (those with good work records).

Table 7

Proportion of Negative Sum of Difference Scores Over Two Administrations of the MMPI--Experimental (Pilot plus Cross-Validation) and Control Groups

	Proportion of Group Having Negative Scores Sum		
	Number In Group	Number Having Negative Sum of Scores	Proportion
Experimental Group			
Working at Follow-up	57	17	.281
Has Worked Since Leaving MRC, But Not Now Working	35	12	.343
Has Never Worked Since Leaving MRC	18	9	.500
No data	31	16	.517
Control Group	29	17	.585

Such results as the above should be taken only as indicative, since the likelihood of their being out of "truly different" distributions of score differences is not exceedingly great. Nevertheless, they are at least suggestive for further research (which should be conducted at the MMPI item level, rather than the scale level, to avoid some of the problems associated with finding useful linear or nonlinear relationships when only scale scores are available).

The test of these scores by the predictive criterion shows again that scores as defined here are only of weak interest. Using these scores to predict post-rehabilitation work success for the experimental group yielded a hit ratio of slightly better than 55 percent, or slightly better than

base rate prediction (lambda = .10). If the assumptions of the random model were acceptable, we would have a probability of getting such a value about 25 percent of the time, if the sample lambda came from a population lambda equal to zero (i.e., the measure adds no information). Therefore, this measure, as it is presently defined, has little promise.

(2) Configural Score Results. The use of configural scoring produced the most useful MMPI-related result of this study. Using the configural scoring technique, a number of combinations of four MMPI scale scores were found for which the hit rate was 70-75 percent (i.e., if the outcomes on each pattern of response were to hold good over repeated applications of the rule, we could use the pattern outcomes to predict work outcome as based on MMPI measured improvement outcome, and be correct in our predictions 70-75 percent of the time). This implies a sample lambda in the .40-.50 range. The further improvement in prediction resulting from use of five-variable and six-variable combinations was considered so small as to not be worth the increase in complexity of interpreting and using results.

Unfortunately, even four-variable patterns are rather unwieldy, when each of the variables is scored four ways (i.e., for each variable the scoring is either, "0" first test less than T=70, second test greater than or equal to T=70; "1" first and second test both greater than or equal to T=70; "2" both first and second tests less than T=70; and, "3" first test greater than or equal to T=70 and second test less than T=70), with $(4)^4 = 256$ possible score patterns involved.

With a sample size no larger than N=82 in the pilot group (the number in the pilot group for whom complete information was available), there will be a large number of empty patterns, and thus an unstable relationship. Therefore it would be useful to find a decision rule which will satisfy three criteria: (a) it will have a high lambda value; (b) it will be stable over a number of replications (that is, whatever decision rule we derive for predicting whether someone will work or not, given that he has a certain improvement pattern, the rule must behave in approximately the same way each time we use it); (c) the rule should be usable by the worker who is responsible for interpreting the test and should, therefore, be simple enough to use in a variety of situations.

In seeking such useful measures of MMPI-related change, we have been successful in deriving rules satisfying the first and third criteria, but not the second. We have

taken some of the most complex rules, which give us a hit rate in the 70-75 percent range, and derived simpler rules, which give us a hit rate in the 65-70 percent range. Unfortunately, cross-validation of these rules with our replication group results in a hit-rate no better than chance or base-rate prediction.

One of these rules is the following: Consider four scales, K, 1 (Hypochondriasis), 7 (Psychasthenia), and 9 (Hypomania) in combination. Then (a) if any of these is coded "0" for a client, predict that the client will not be working at follow-up; (b) for the remaining subjects, predict that all of those having a "3" coded on any scale will be found working at follow-up; (c) of those subjects now remaining, predict that all subjects with a "2" in the Hypochondriasis scale, and a "1" in at least one other scale, will be found working at follow-up; (d) for the remainder (the majority of which will usually be those who, on all four scales, began under 70 and ended there), predict that they will not be found working at follow-up. Using a rule of this simplicity, a hit rate of better than 70 percent can be achieved.

If we denote each separate prediction group as (a) - (d), the details of this result are as follows:

Table 8(a)

An MMPI Follow-up - Employment Prediction Rule, Using Scales K, 1, 7, and 9

Prediction Group	Prediction	Actual Outcome at Follow-up	
		Working	Not Working
(a) All with '0' on at least one scale	"Won't Work"	2	10
(b) All remaining with a '3' on at least one scale	"Will Work"	16	3
(c) All remaining with a '2' on Hypochondriasis and at least one scale scored '1'	"Will Work"	6	2
(d) All remaining	"Won't Work"	17	26
Total		41	41

Such a decision rule results in two kinds of correct predictions and two kinds of errors, as follows:

Table 8(b)

Results of Use of an MMPI Employment-Prediction Rule

	Predicted Will Work	Predicted Won't Work	Total
Actually Working at Follow-up	22	19	41
Actually Not Working at Follow-up	5	36	41
Total	27	55	82

The two kinds of correct prediction (predicted working at follow-up; found working at follow-up, n=22; predicted not working at follow-up; found not working at follow-up, n=36, total correct is 58 out of 82, or 70 percent) result in much improvement over the base-rate prediction level.

However, when this decision rule was cross-validated, only 14 of the 28 in the cross-validation group were correctly classified, or roughly about what we could expect by using chance or base-rate prediction (i.e., lambda = 0.00). The small number in the cross-validation group for whom complete data were avilable (n=28) might have something to do with this; but the rule, if it is a good one, should operate well with smaller groups; and, under a random sampling model, the result we obtained should happen only rarely, if the decision rule were a true discriminator. With such mixed evidence (a high lambda on the original rule, but a zero lambda on the cross-validation), it would seem that such rules should be investigated carefully for possible improvements - especially in the cases for whom the rule creates errors in decision, but should not be used in any practical application without much further research. A more detailed discussion of reasons for deterioration of results under cross-validation follows the next section.

(3) Using Other Information: A natural accompaniment to questions of prediction through measures of change under rehabilitation is the question of the efficiency of information about various states of the client at entry into rehabilitation as predictors of outcome. In a sense, such questions have less theoretical significance, since they are not attempting to find relationships in which psychological theory concerning the relationship between improvement under treatment and later outcome is the issue at hand. Here we are simply interested in the relationship between the client's state at entry and his later work state. With such measures, we are not adding information about the response to treatment, or even assuming

that there is one, no assumptions are made about the effects of treatment.

The following variables were coded: sex, entry score on each scale of the MMPI (less than T=70 on a scale versus T=70 or greater on a scale), marital status (single, separated or divorced, versus widowed or married), number of weeks in program (less than five weeks versus five weeks or more), IQ (less than 100 versus 100 or greater), age (less than median age for group versus greater than median age for group; the pilot group median was 24 years, the replication group median 31 years, and the control group median 31 years), medical status (see Appendix V for material on this coding).

Because only dichotomous predictors to a dichotomous criterion were involved, dealing with these variables involved less complexity than the quartotomous variable defined earlier for a measure of change. Here, even a six variable combination of dichotomous predictors involved far fewer patterns (64) than a four-variable combination of quartotomous predictors (256 patterns).

In looking for sets of best predictor combinations, we examined demographic and medical variables separately from the MMPI variables, then examined all variables together to find best predictor combinations. It was notable that MMPI variables, taken alone, are lowest in predictive power. The common demographic variables rank highest, with the psychological and medical variables ranking second.

The best and second-best combinations of one, two,..., five or six variables are listed together, by hit rate of the best combination: (See following page)

It is interesting that even one variable gives us considerable improvement (lambda = .275) in information for the prediction of outcome. Here, our prediction would be simple: if female, predict that the client will not be working at follow-up; if male, predict that the client will be working at follow-up (the hit rate improves from .51 to .65). After the first variable the rate of increase becomes almost constant for each additional variable added (about three percent additional predictive efficiency per additional variable, for the non-MMPI variables). The addition of the MMPI variables in combination provides a considerable increase in predictive efficiency. For example, the combination of five MMPI, medical, and demographic variables provide about 3.5 percent higher predictive efficiency than the combination of five non-MMPI variables (.805 versus .768). The MMPI variables taken alone were rather inefficient; the best six-variable-combination provided no more predictive efficiency than the best four-variable medical-demographic combination.

Table 9

Best Predictor Combination Related to the Outcome
(Employed at Follow-up After Receipt of MRC Services)
By Number and Type of Predictor Variables Used

Number of Variables	Hit Rate	Best Variable Combination	Second-Best Combination
1 (without MMPI)	.646	Sex	Marital Status
2 (without MMPI)	.671	Sex, IQ	Sex, Marital Status
3 (without MMPI)	.707	Sex, Age, Marital Status	Sex, Congenitial Medical Disability, IQ
4 (without MMPI)	.732	Sex, Motor Disability Psychiatric Disability, Age	Sex, Psychiatric Disability, Sensory Di-ability, IQ
5 (without MMPI)	.768	Sex, Motor Disability Psychiatric Disability, Age, Weeks in Program	Sex, Motor Disability, Psychiatric Disability Age, IQ
5 (MMPI & Other)	.805	Sex, Age, Marital Status, and MMPI Scales 2, and 8	Orthopedic Disability, Sensory Disability, Marital Status, and MMPI Scales 4, and 7
6 (without MMPI)	.817	Sex, Congenital Disability, Motor Disability, Age, IQ Weeks in Program	Not Calculated
6 (MMPI alone)	.732	MMPI Scales K,1,2,3,7,9	MMPI Scales ES,K,1,3,7,9

As with the measures of psychological change, a decision rule using these combinations should meet two basic requirements in order to be of further interest: (a) it must be amenable to wide use, and (b) it must hold up under cross-validation.

With dichotomous coding the use of one variable implies a prediction for each of two groups; two variables imply predictions for four groups, etc. We felt that any finer delineation of such small samples would not yield rules with any probability of standing up over repeated use. As an example of how just one variable may give excellent improvement over base rate, with the addition of another variable giving further improvement (but at decreased rate of improvement), consider the use of sex and IQ in prediction of outcome.

If we consider sex alone, we derive the prediction table:

Client Type	Working Number	Not Working Number	Percent Working
Male	29	16	64
Female	13	24	35

Using the prediction rule - if male, predict work; if female, predict no work - we arrive at a hit rate of .65.

If we now add another variable, IQ, we get a (slightly improved) hit rate of .67.

Client Type	Working Number	Not Working Number	Percent Working
High-IQ Females	6	4	60
Low-IQ Females	7	20	26
High-IQ Males	12	6	67
Low-IQ Males	17	10	63

The variables discussed above were not the only variable combinations which gave some indication of predictive power. The high-MMPI result, as previously noted, produced a high hit rate (.732), with a consequent excellent lambda value (.45). However, this was a six-variable result (K, 1,2,3,7,9). A simplification of this result produced a rule which utilized only three of the variables (K,3,7). The rule: For all clients having a K-scale score equal to or greater than T=70, or a K-scale score of less than 70 and only one of the other two equal to or greater than 70, predict that they will be found working at follow-up; for all other clients, predict that they will not be found working at follow-up. The resultant is a hit rate of .67, a not intolerable loss, given the simplification.

There were also simplifiable combinations of medical, demographic, and MMPI variables. A five-variable result, which gave a hit rate of .805 using (a) orthopedic handicap involved or not; (b) sensory handicap involved or not; (c) single versus ever married; (d) under T=70 versus 70 or over on MMPI Psychopathic Deviate Scale; and (e) under T=70 versus T=70 or over on the MMPI Psychasthenia Scale was collapsed into the simplest of decision rules: If the

subject had neither of the handicaps, was single, and had a T-score of less than 70 on both scales, predict that the subject would not be found working at follow-up; otherwise, predict that the subject would be working at follow-up. The result was a hit rate of .72, the best result of all "collapsed decision rules" defined for the study.

The result of the prediction was the following:

	Predicted Work at Follow-up	Predicted No Work at Follow-up	Total
Actually Working at Follow-up	30	12	42
Actually Not Working at Follow-up	11	29	40
Total	41	41	82

In the cross-validation of the last three simplified rules, the results were again no better than chance or base-rate prediction, indicating that these rules also add nothing to our discriminative or predictive power.

On the Deterioration of Predictor - Equation Results: The general deterioration of results which we have experienced in using these rules could be ascribed to one or a combination of reasons: for example, our cross-validation group might not "really" be from the same population. As we noted before, those chosen for each group do not satisfy any random-selection model. Since the pilot and replication groups were taken at two differing time intervals, it is quite possible that the population of those being served had changed in some basic fashion (it is known, for example, that different economic conditions, which induce differing degrees of unemployment, result in somewhat different groups appearing at MRC from the same agency). As an indication of this possibility, the median age shifts considerably, as between pilot and replication group clients who reported outcome data (from 24 to 31). Thus, the "young male" group of the pilot group, which contributed heavily to the proportion employed at follow-up for the pilot group (more than two-thirds of this group, which amounted to one-fourth of the pilot group, were employed at follow-up is non-existent in the cross-validation group. Yet, the cross-validation group achieved the same proportion of the total group employed at follow-up as the pilot group (about one-half).

Lastly, and probably most relevant to the matter, the search technique used to find "best predictive combinations" of variables is a computerized form of a practice first noted by R. A. Fisher in his famous paragraph 24 of The Design of Experiments (pp. 57-58, 5th Edition, 1949), and the subject

of a number of well-known solutions since (e.g., Scheffe's a posteriori technique). The problem was first discussed within the context of the randomized blocks design, in which means for more than two treatments were being compared. It is worth quoting:

> "Comparisons, which the experiment was designed to make, may, of course, be made without hesitation. It is comparisons suggested subsequently, by a scrutiny of the results themselves, that are open to suspicion; for if the variants are numerous, a comparison of the highest with the lowest observed value, picked out from the results, will often appear to be significant, even from undifferentiated material. Properly, such unforeseen effects should be regarded only as suggestions for future experimentation, in which they can be deliberately tested. To form a preliminary opinion as to the strength of the evidence, it is sometimes useful to consider how many similar comparisons would have been from the start equally plausible. Thus, in comparing the best with the worst of ten tested varieties, we have chosen the pair with the largest apparent difference out of 45 pairs, which might equally have been chosen. We might, therefore, require the probability of the observed difference to be as small as 1 in 900, instead of 1 in 20, before attaching statistical significance to the contrast."

Analogously, in order to find "best predictor combinations" of variables, we have, in the case of the MMPI change measures, searched 1,001 combinations; in the case of the MMPI entry scores, taken alone, we searched 3,003 combinations; for the demographic and medical variables, taken alone, we searched 210 combinations; and, for the MMPI entry scores plus the demographic and medical variables, we searched 42,504 combinations. In order to achieve an actual .05 significance probability, we would need to utilize (in the absence of the Scheffe results) a nominal significance probability roughly on the order of p less than .000001 (if we had satisfied the probability assumptions in the way we chose our original group). This provides the rationale for the cross-validation group, since it serves as the "future experimentation" group, in which these "unforeseen effects...can be deliberately tested." To the extent that we can "trust" the cross-validation group, then, we can accept the negative evidence supplied by it, in trying out the decision rules derived from the many-combination search of the pilot group.

Such deterioration in predictive power indicates that we have not yet found the Rosetta Stone in discriminating successful from unsuccessful at the point of selection for vocational rehabilitation or even after MMPI measurements are made on the client's change during the rehabilitation process.

In the other two published "success prediction" projects in which vocational rehabilitation outcome of MMPI predictors were used, not much greater success resulted. The more successful was that of Gough, Wenk, and Rozynko, in which a lambda of .16 (our calculation from the author's data) was achieved under cross-validation, using a combination of the California Base Expectancy Scale and the California Personality Inventory (a combination of these two plus the MMPI was approximately as powerful; i.e., the MMPI added nothing here) to predict success or failure in parole outcome. In the other pub-

lished study, DeMann developed an eight-variable equation for prediction of success in vocational rehabilitation outcome. In cross-validation, using a group split 84:16 on the rehabilitated-not rehabilitated outcome measure, the equation predicted only 65 percent correctly, or far less than base rate prediction.

Such results as were found here and in the other studies cited would seem to indicate that, if useful predictive equations are to be derived, such that theory is also satisfied, we must turn to the operational definition of substantive variables, as opposed to the "available-demographic" or test variables which are usually those investigated (as was the case here). That is, further investigation should concern itself with those vocational rehabilitation workers who are themselves "good predictors" of client success, and an attempt should be made to derive a heuristic programming of the decision rules that these workers are using in their operations. Such a simulation-model approach would seem, given its success in other areas, the most promising one available at the present time.

V. SIGNIFICANCE OF THE RESULTS OF THIS RESEARCH

Every vocational rehabilitation center staff faces the following basic problem: A constant stream of clients come to it, some of whom it can help very much, some of whom it can help somewhat, and some of whom it cannot help at all.

The problem for the center staff is to discriminate between these groups so that it can - under all conditions of demand for its services - refer elsewhere those potential clients whose treatment would be a waste of time and the public's money. For example, it is not useful for the center to accept borderline referrals who can probably do quite well occupationally without the help of the center, or those potential clients who are at a level such that the services of the center's staff will not help the client at all.

Under most conditions, it is also useful to be able to distinguish - among those accepted for services - those who will benefit less from the staff's services from those who will benefit more from such services.

There are at least two reasons for wanting to do this. One is that the agency, once it has performed its "differential diagnosis," may be able to allocate potential clients into different programs. For example, those who would benefit from the agency's services, but have a higher probability of job success, can go directly into the agency's "intensive services." Some, however, of the lower probability-of-success group will need a short period of preparatory work before they will be able to take full advantage of the agency's intensive services. There might also exist a group with even lower probabilities of success, who would need rather long-term preparatory services before they could enter the intensive vocational services.

Secondly, given the particular social commitment of the agency, based upon its ideas about what subgroup it can work best with, or what the relative benefits are to society in dealing with the higher and the lower probability-of-success groups, the agency will change its admissions policies somewhat under different conditions of demand for its (presumably limited amount of) services.

For example, if the agency had decided that it could do its best job with the higher and medium probability-of-return-to-work groups, it would then wish to screen out the lower probability group for referral elsewhere, or for delay in rendering vocational services until demand slackened. Or if, contrariwise, the agency had decided that its best work was done with those having low probability of employment, and it believed that high demand for its services was positively correlated with labor shortages, it might focus on taking low and medium-probability-of-success group and filtering out the higher probability applicants for services, on the ground that, in a time of labor shortage, many employers are more than happy to serve as "vocational rehabilitators" for these "easier" cases. Whatever the conditions of demand for service or the agency's social policy, then, it has the need to distinguish between types of potential clients, according to the poten-

tial benefit the agency's services will have for them. A tabulation of what might be a vocational rehabilitation agency's possible actions in its admissions-decision policies, given these three kinds of determinants, is given below

Table 11

The Possible Admissions Screening Actions of a Vocational Rehabilitation Center, Given Its Social Policy, Demand for Its Services, and the Probability of the Potential Client's Returning to Work

	SOCIAL POLICY OF AGENCY			
	Take Those Most Likely to Succeed, If There Is a Choice		Take Those Who Need Services Most, If There Is a Choice	
	DEMAND CONDITION		DEMAND CONDITION	
PROBABILITY OF CLIENT'S RETURNING TO WORK	Great Demand	Slack Demand	Great Demand	Slack Demand
Work a Certainty without Agency's Services	Do Not Accept	Do Not Accept	Do Not Accept	Do Not Accept
High Probability of Work with Agency Services	Send Straight to Intensive Team Services	Send Straight To Team Services	Do Not Accept	Send Straight to Intensive Team Services
Medium Probability of work with Agency Services	Send to a Short Preparatory Course Prior to Entry Into Team Services	Send to a Short Preparatory Course Prior to Entry Into Team Services	Send to a Short Preparatory Course Prior to Entry Into Team Services	Send to a Short Preparatory Course Prior to Entry Into Team Services
Low Probability of Work with Agency Services	Do Not Accept	Send to Exhaustive Preparatory Course	Send to Exhaustive Preparatory Course	Send to Exhaustive Preparatory Course
No Work a Certainty Even with Agency Services	Do Not Accept	Do Not Accept	Do Not Accept	Do Not Accept

We have discussed the kinds of choices to be made and some determinants of these choices. Now, we need some methods of making such choices. That is, we need to know how to recognize which potential clients should go into each path. If we are to allocate potential clients into each of these paths in some efficient way, we are really making a prediction about whether they will work or not and under which conditions of rehabilitation they will have their best chance of returning to work. Therefore, if we have some method of predicting the potential client's probability of returning to work, given the agency services, and can do this with entry data alone, we have a method for allocating the potential clients to their proper paths. The attempted development of such a method was discussed as the last portion of the technical report.

We would also like to know how well various clients have done during rehabilitation, so that we can predict, after making some observations or measurements of their progress in the agency's programs, their probability of returning to work. A method for making such measurements would permit us to know which clients to keep on for further services (because they have not yet "responded to treatment") and which ones to let go. In this way, we can make more optimal use of our facilities, not keeping on those who show themselves ready to leave and keeping on those who have shown themselves not yet ready to leave.

It is to the study of this problem that the main portion of this research was directed. Since it was believed that a major component of vocational disability and a major component of positive changes reside in the personality dynamics of the client, the chief effort of the study was devoted to examining possible measures of client change under treatment, based upon MMPI data, in order to develop such a method. It was also recognized that the development of a successful measurement method would have important implications for MMPI theory.

The design of the study insured that the critical decision on whether we had found bases for predicting success in the two different cases - prediction of post-rehabilitation success given what we know of the client at entry, and prediction of post-rehabilitation success given a measure of change under treatment which we could observe after some time in treatment - would be made on a conservative basis. The results were mixed.

In both cases, types of information were defined which gave considerable improvement in predictive power. However, in no case was there a successful cross-validation. Since cross-validation results are crucial under the assumptions and procedures used in this study, the major results are negative. It should be noted, however, that problems of construction of each of the groups used here (pilot, control, and cross-validation groups) may be responsible for the negative results. Further, as noted in the technical discussion, there were difficulties in construction of the criterion of "success" which may also have contributed to this result. The major effect of this study, in this respect, therefore is to provide the basis for a well-designed study which could produce the theoretical and practical results, the need for which was clarified by this study.

Aside from the main results of the study, certain valuable secondary results were produced. Among these:

1) One of the major information problems in vocational rehabilitation arises from the fact that the majority of applicants for vocational rehabilitation are multi-problem cases. However, most information in the field concerns only the "primary problem" of the client. A multiple-problem coding system, with rules for coding, was devised as part of this study, and found to be relatively successful.*

2) Although the "base-rate prediction trap" and methods for its avoidance have been known to the literature of prediction in general and vocational psychological matters for some years (dating from Paul Meehl's article in the mid-fifties), it seems to have been honored more in the breach, in research reports where it was a relevant consideration, than in its application (see, for example, the two reports of research on prediction discussed in the technical report of this study). The problem is defined in operational terms, with examples, here. A simple statistical method for solving the problem is also given, using Guttmann's lambda, which is a measure of the decrease in the probability of predicting incorrectly, over and above that given by knowledge of the base rate of occurrence of a phenomenon, as we add pieces of information. Although the associated distribution theory was used here in only one example (because of the underlying probabilistic problems involved in the way the study groups were constructed), confidence interval methods for this statistic, derived by means of asymptotic theory, and applicable on a fairly small sample basis are available for application in problems where they are relevant. Therefore, in any study satisfying design assumptions, we may compare prediction methods for relative predictive power, decide whether the predictive power of an equation developed through these methods is "significantly different from zero", and the like.*

3) Some methods of definition of a "measure of psychological change" during vocational rehabilitation were explored, and a promising form of configural scoring was developed. As was noted in the technical report, the possible number of such measures which may develop is extremely large. However, the results obtained here pointed to some further interesting possibilities using scale score results. Future research in this area should be planned to take advantage of known "common configurations" and the personality information which is generally considered by MMPI theorists to be correlated with these configurations. Such research would then allow measures of change which have been validated by a correspondence with an economic criterion to be related to current psychological theory.

4) A major problem in evaluating what goes on during the rehabilitation process has been the lack of a well-defined criterion of success. Some of the complexities inherent in the use of a criterion which involves only the judgment that someone is working at time of follow-up (but not necessarily

---

*(Note: Information on these methods, and computer programs which simplify their use, is available from the project's biostatistician, Mr. Copeland)

much before or much after; i.e., one of the problems in "success" is that some of its components are stochastic) were examined, so that what is involved in using a criterion at the low end of the "complexity spectrum" might be somewhat better understood. The problem is difficult enough so that the problem of constructing a criterion deserves "on-site" research, in which the post-rehabilitation progress of the client is closely monitored, so that a quantitative criterion measure can be constructed on a large-sample basis. Only then can we accurately assess the predictive relationships between entry information, "in-process" information, and the criterion. And only when we have done this can we derive useful entry-predictors-of-success and useful measures of change under rehabilitation.

BIBLIOGRAPHY

1. Report of Research, November 30, 1963, Minneapolis Rehabilitation Center, Inc.

2. Welsh, G.S., and Dahlstrom, W.V., Basic Readings on the MMPI in Psychology and Medicine, University of Minnesota Press, Minneapolis, 1956.

3. Goodman, L. and Kruskal, W. H., Measures of Association for Cross Classifications, Journal of the American Statistical Association, pp. 732-764, December, 1954.

4. Cronbach, L. J., and Gleser, G.G., Psychological Tests and Personnel Decisions, 2nd Ed., University of Illinois Press, 1965.

5. Fisher, R.A., The Design of Experiments, 5th Edition, Hafner, 1949.

6. Gough, H.G., Wenk, E.A., Rozynko, V.V., Parole Outcome as Predicted From the CPI, the MMPI, and a Base Expectancy Table, Journal of Abnormal Psychology, 1965, Volume 70, No. 6, pp. 432-441.

7. DeMann, M.M., A Predictive Study of Rehabilitation Counseling Outcomes, Journal of Consulting Psychology, Vol. 10, No. 4, 1963.

APPENDIX I

TABLES AND DISCUSSION OF CHARACTERISTICS OF EXPERIMENTAL (PILOT AND REPLICATION) AND CONTROL GROUPS

In a number of other tables in this report, we will report data concerning the subjects in the bivariate or trivariate modes, as well as the more usual univariate mode. Because the data is presented in this fashion, far more information is more available for any inspection of subgroup data that the reader may care to make than is possible in the usual mode of presentation in which all data is presented only in the univariate mode.

1. Age, Sex, and Education of Experimental and Control Groups. It will be noted that the data for the 101 subjects in Experimental Group I (the Pilot Group) and the 40 subjects of Experimental Group II (the Replication Group) have been grouped together for these comparisons (see Table 1).

Although the sex ratio for the experimental groups is fairly even (about 56:44), fewer than a quarter of the control group (7 of 29) were female. At the same time, the average age for the control group is somewhat greater (slightly more than a fifth of the experimentals are under 20 years of age, but only 1 of 29 controls). While the control group distribution of years of education was somewhat better than that of the experimental group (one-half of the experimentals had 12 years or better, versus about two-thirds of the control group with equal attainment), the sex differences within groups were somewhat greater. Males in the experimental groups were the worst in this regard.

2. Age and I.Q. The distribution of IQ is about equivalent for control and experimental groups (see Table 2). Whether there is an age-IQ interaction in the control group is not really decidable--but the existence of such an interaction for the experimental group is clear. Those accepted for services prior to the age of 40 are clearly a low-IQ group (about 80 percent of those under 20, about 70 percent of those between 20 and 39, but only about 30 percent of those 40 and over have IQ scores of less than 100). The effects of this age-IQ interaction on work after rehabilitation services will become evident later in this section.

3. Marital Status, by Group. A last demographic comparison between the two groups illuminates one especially interesting characteristic of the groups entering MRC--the heavy preponderance of single versus ever-married among the clients. Here, those classified as single amounted to about two-thirds of the experimental group (98 of 141) and one-half of the control group (14 of 29). Only about one-sixth of the experimental group (25 of 141) and one-fourth of the control group (7 of 29) were listed as married and currently living with their spouses.

Table I

Education, Age, and Sex, by Subject Group

AGE	MALES - EXPER. GROUP				
	5-8 yrs	9-11 yrs	12 yrs	13+ yrs	Total
15-19	1	7	5	-	13
20-24	3	6	10	1	20
25-39	8	10	6	2	26
40+	4	5	7	4	20
Total	16	28	28	7	79

AGE	FEMALES - EXPER. GROUP				
	5-8 yrs	9-11 yrs	12 yrs	13+ yrs	Total
15-19	-	4	12	1	17
20-24	1	3	12	2	18
25-39	1	1	7	4	13
40+	5	1	4	4	14
Total	7	9	35	11	62

AGE	MALES - CONTROL GROUP				
	5-8 yrs	9-11 yrs	12 yrs	13+ yrs	Total
15-19	-	-	1	-	1
20-24	-	3	5	1	9
25-39	1	1	3	-	5
40+	1	2	4	-	7
Total	2	6	13	1	22

AGE	FEMALES - CONTROL GROUP				
	5-8 yrs	9-11 yrs	12 yrs	13+ yrs	Total
15-19	-	-	-	-	-
20-24	-	-	-	-	-
25-39	-	1	3	2	6
40+	-	-	1	-	1
Total	-	1	4	2	7

Table 1 (Continued)

Education, Age, and Sex, by Subject Group

AGE	TOTAL EXPER. GROUP				
	5-8 yrs	9-11 yrs	12 yrs	13+ yrs	Total
15-19	1	11	17	1	30
20-24	4	9	22	3	38
25-39	9	11	13	6	39
40+	9	6	11	8	34
Total	23	37	63	18	141

AGE	TOTAL CONTROL GROUP				
	5-8 yrs	9-11 yrs	12 yrs	13+ yrs	Total
15-19	-	-	1	-	1
20-24	-	3	5	1	9
25-39	1	2	6	2	11
40+	1	2	5	-	8
Total	2	7	17	3	29

Table 2

Age and IQ, by Group

CONTROL GROUP					
	IQ				
AGE	60-89	90-99	100-109	110-129	Total
15-19	-	-	1	-	1
20-24	2	4	3	-	9
25-39	4	3	3	1	11
40+	3	2	-	3	8
Total	9	9	7	4	29

TOTAL EXPER. GROUP					
	IQ				
AGE	60-89	90-99	100-109	110-129	Total
15-19	8	16	5	1	30
20-24	17	10	6	5	38
25-39	12	14	8	5	39
40+	2	8	12	12	34
Total	39	48	31	23	141

Table 3

Marital Status, By Group

Marital Status	Group	
	Experimental	Control
Single	98	14
Married	25	7
Separated	6	2
Divorced	11	4
Widowed	1	2
Total	141	29

APPENDIX II

MEANS AND STANDARD DEVIATIONS OF FOURTEEN MMPI SCALE SCORES (T-SCORES), TAKEN OVER TWO OCCASIONS OF TESTING, FOR THREE GROUPS (PILOT GROUP, REPLICATION GROUP, CONTROL GROUP), BY SEX

Table 1

Cumulative Distribution of Sum of Difference Scores (Σxi) on Two Administrations of the MMPI, for Three Vocational Rehabilitation Groups

Sum-of-Difference-Score (Σxi) Interval	Pilot Group			Cross-Validation Group			Control Group		
	fj	pj	cpj	fj	pj	cpj	fj	pj	cpj
≤ - 120	0	0	0	0	0	0	2	6.85	6.85
-119 - -100	1	.99	.99	1	2.5	2.5	0	0	6.85
-99 - -80	4	2.96	4.95	2	5.0	7.5	1	3.45	10.30
-79 - -60	1	.99	5.94	0	0	7.5	3	10.35	20.65
-59 - -40	6	5.94	11.88	2	5.0	12.5	2	6.85	27.50
-39 - -20	7	6.93	18.81	5	12.5	25.0	3	10.35	37.85
-19 - 0	14	13.86	32.67	10	25.0	50.0	4	13.80	51.65
1 - 20	20	19.80	52.47	6	15.0	65.0	5	17.25	68.90
21 - 40	20	19.80	72.27	5	12.5	77.5	3	10.35	79.25
41 - 60	13	12.87	85.14	3	7.5	85.0	3	10.35	89.60
61 - 80	8	7.92	93.06	4	10.0	95.0	0	0	89.60
81 - 100	4	3.96	97.02	1	2.5	97.5	1	3.45	93.05
101 - 120	2	1.98	99.00	1	2.5	100.0	2	6.85	10.00
121 - 140	1	.99	99.99	0	0	100.0	0	-	-
Total	101	100		40	100		29	100	100

Note: fj = the frequency of sum-of-difference scores in a group which are in a given sum-of-difference scores interval, j;
pj = the percentage of scores, in a group, which are in a given sum-of-difference-scores interval, j;
cpj = cumulative percentage of Σxi Scores, i.e., cpj = the percentage of scores which are less than the least score in interval j + 1.

Table 2

Cumulative Distribution of Sum of Squared Difference Scores (xi^2) on Two Administrations of the MMPI, for Three Vocational Rehabilitation Groups

Sum-of-Squared Difference Scores (x_i^2) Interval	Pilot Group			Cross-Validation Group			Control Group		
	fj	pj	cpj	fj	pj	cpj	fj	pj	cpj
- 200	4	3.96	3.96	4	10	10	-	-	-
200 - 299	6	5.94	9.90	1	2.5	12.5	4	13.80	13.80
300 - 399	10	9.90	19.80	5	12.5	25	1	3.45	17.25
400 - 499	11	10.89	30.69	2	5.0	30	5	17.25	34.50
500 - 599	8	7.92	38.61	8	20	50	1	3.45	37.95
600 - 699	10	9.90	48.51	2	5.0	55	3	10.35	48.30
700 - 799	8	7.92	56.43	4	10.0	65	1	3.45	51.75
800 - 899	4	3.96	60.39	3	7.5	72.5	3	10.35	62.10
900 - 999	10	9.90	70.29	3	7.5	80.0	2	6.85	68.95
1000 - 1099	7	6.93	77.22	1	2.5	82.5	3	10.35	79.30
1100 - 1199	4	3.96	81.18	2	5.0	87.5	-	-	-
1200 - 1299	3	2.97	84.15	-	-	-	1	3.45	82.75
1300 - 1399	1	.99	85.14	1	2.5	90.0	-	-	-
1400 - 1999	9	8.92	94.06	3	7.5	97.5	-	-	-
- 2000	6	5.94	100.00	1	2.5	100.0	5	17.25	100.00
Total	101	100		40	100		29	100	

Note: pj = percentage of scores in group which are in a given sum-of-squared-difference scores interval, j;
cpj = cumulative percentage of xi^2 scores, i.e., cpj = the percentage of scores which are less than the least score in interval j + 1.

Table 3

Pilot Group

Scale	Male (N=57)				Female (N=44)			
	Initial Score		Final Score		Initial Score		Final Score	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
ES	50.49	11.06	51.05	11.47	48.61	7.70	52.30	9.11
L	53.77	8.80	53.33	8.60	53.75	8.15	55.07	8.93
F	58.65	12.85	57.54	12.21	56.41	7.87	55.68	9.23
K	55.07	10.75	55.51	11.13	53.32	8.38	55.80	9.67
1	58.42	14.36	56.67	13.50	55.36	12.66	52.36	12.68
2	62.42	13.53	59.61	13.40	60.89	11.50	58.68	12.24
3	59.89	11.95	59.49	10.30	56.55	10.70	55.14	11.36
4	64.40	11.40	62.35	10.77	59.34	11.25	61.86	11.29
5	56.44	10.34	55.04	9.83	52.14	9.21	50.14	9.21
6	57.95	10.94	56.79	10.53	57.45	8.78	55.66	10.64
7	60.67	10.70	58.60	11.74	58.57	9.05	56.57	8.63
8	62.93	13.52	62.74	14.28	60.32	9.79	59.11	10.62
9	57.42	12.11	57.77	11.67	56.95	10.35	57.20	11.99
0	52.84	10.19	52.77	10.18	60.50	9.93	57.25	10.35

Table 4

Cross-Validation

Scale	Male (N=22)				Female (N=18)			
	Initial Score Mean	S.D.	Final Score Mean	S.D.	Initial Score Mean	S.D.	Final Score Mean	S.D.
ES	51.18	12.40	53.36	11.54	52.72	10.45	53.89	10.31
L	51.50	7.87	52.45	7.54	54.28	11.49	56.33	12.90
F	57.14	11.91	56.09	11.76	58.83	11.64	59.28	14.29
K	56.68	9.02	58.95	9.21	53.94	11.31	55.44	11.34
1	59.95	12.03	59.18	11.47	49.33	8.45	52.28	10.05
2.	66.95	9.71	62.27	9.03	60.61	8.25	58.11	8.12
3.	61.55	8.05	60.77	8.31	54.06	10.10	55.00	12.48
4.	65.23	11.00	63.95	10.23	63.17	12.35	63.06	13.14
5.	54.95	9.41	55.59	11.00	51.56	12.94	52.72	14.64
6.	56.27	10.89	56.00	12.96	57.28	9.72	58.67	12.67
7.	61.68	9.59	61.95	10.76	59.83	9.29	58.00	10.09
8.	62.77	14.64	60.86	14.37	64.06	15.04	63.06	14.69
9.	59.00	13.84	59.27	11.87	59.11	8.60	58.67	11.05
0.	52.64	9.55	50.64	10.99	55.17	8.28	55.44	9.17

Table 5

Control Group

Scale	Male (N=22)				Female (N=7)			
	Initial Group Mean	S.D.	Final Score Mean	S.D.	Initial Score Mean	S.D.	Final Score Mean	S.D.
ES	46.77	12.42	48.05	14.80	45.57	11.1	47.00	13.45
L	53.00	8.13	53.95	9.03	52.00	10.63	52.29	12.93
F	59.82	16.54	62.23	18.69	64.43	15.58	62.29	10.59
K	55.23	10.40	58.14	12.24	59.00	11.80	57.71	11.86
1	60.68	10.28	64.41	13.27	53.71	5.94	52.86	5.01
2	63.82	12.48	65.27	15.65	63.71	16.26	67.57	16.68
3	60.86	9.19	64.09	10.38	61.00	6.06	59.86	10.09
4	65.27	11.05	69.14	14.05	69.71	13.38	68.00	14.59
5	57.50	11.37	59.95	10.96	47.29	7.61	49.29	11.70
6	60.68	11.54	62.23	16.23	69.57	12.54	67.14	9.86
7	61.77	11.80	64.59	15.09	63.43	11.12	56.00	17.68
8	65.41	16.56	68.77	17.24	63.57	10.61	68.43	12.82
9	57.05	14.08	62.32	10.92	57.57	14.35	53.14	13.69
0	52.73	8.88	51.00	10.51	55.29	13.97	55.43	12.01

APPENDIX III

EXPERIMENTAL GROUP (PILOT AND CROSS-VALIDATION) OUTCOME DATA, BY AGE, SEX, EDUCATION, IQ, AND MARITAL STATUS

Data are included separately for the Pilot and Cross Validation groups, categorized by outcome (now working versus has worked - but is not now working versus never worked versus no data on follow-up), in the following tables:

1) By Age and Education

2) By Age and IQ

3) By Age and Sex

4) By Marital Status and Sex

5) By Five Collapsed Variables of Interest

Table III-1a

Pilot Group - Outcome Status

Age	Education	Not Working	Not Now Working		No Data On Follow-up	Total
			Has Worked	Never Worked		
14-19	0-9 yrs	4	1	-	1	6
	10-11	-	1	1	1	3
	12	5	3	4	2	14
	13-16	-	-	-	-	-
	Total	9	5	5	4	23
20-23	0-9	1	3	-	2	6
	10-11	1	1	-	-	2
	12	4	6	4	2	16
	13-16	-	-	-	1	1
	Total	6	10	4	5	25
24-33	0-9	2	-	-	-	2
	10-11	5	2	-	2	9
	12	8	2	-	1	11
	13-16	1	1	-	-	2
	Total	16	5	0	3	24
34-63	0-9	4	1	2	4	11
	10-11	3	2	-	-	5
	12	2	2	3	1	8
	13-16	2	1	-	2	5
	Total	11	6	5	7	29
Total		42	26	14	19	101

Table III-1b

Cross-Validation Group - Outcome Status

Age	Education	Not Now Working: Now Working	Has Worked	Never Worked	No Data On Follow-up	Total
14-20	0-9 yrs	-	-	-	2	2
	10-11	-	-	-	1	1
	12	3	-	-	-	3
	13-16	-	1	-	-	1
	Total	3	1	0	3	7
21-30	0-9	1	-	-	1	2
	10-11	1	-	-	-	1
	12	1	-	-	4	5
	13-16	2	2	-	1	5
	Total	5	2	0	6	13
31-40	0-9	-	2	-	1	3
	10-11	2	-	-	1	3
	12	-	2	-	1	3
	13-16	1	-	-	-	1
	Total	3	4	0	3	10
41-65	0-9	2	1	1	-	4
	10-11	-	-	-	-	0
	12	1	-	2	-	3
	13-16	1	1	1	-	3
	Total	4	2	4	0	10
Total		15	9	4	12	40

Table III-2a

Pilot Group - Outcome Status

Age	IQ	Not Working	Not Now Working		No Data On Follow-Up	Total
			Has Worked	Never Worked		
14-19	60-89	1	-	1	2	4
	90-99	5	5	2	2	14
	100-109	3	-	2	-	5
	110-129	-	-	-	-	-
	Total	9	5	5	4	23
20-23	60-89	4	6	2	3	15
	90-99	-	3	1	-	4
	100-109	1	-	-	2	3
	110-129	1	1	1	-	3
	Total	6	10	4	5	25
24-33	60-89	6	1	-	-	7
	90-99	3	3	-	3	9
	100-109	6	-	-	-	6
	110-129	1	1	-	-	2
	Total	16	5	0	3	24
34-63	60-89±	-	2	2	-	4
	90-99	5	1	1	3	10
	100-109	4	1	1	3	9
	110-129	2	2	1	1	6
	Total	11	6	5	7	29
Total		42	26	14	19	101

Table III-2b

Cross-Validation Group - Outcome Status

Age	IQ	Now Working	Not Now Working		No Data On Follow-up	Total
			Has Worked	Never Worked		
14-20	60-89	1	-	-	3	4
	90-99	2	-	-	-	2
	100-109	-	-	-	-	0
	110-129	-	1	-	-	1
	Total	3	1	0	3	7
21-30	60-89	-	-	-	1	1
	90-99	2	1	-	3	6
	100-109	3	-	-	-	3
	110-129	-	1	-	2	3
	Total	5	2	0	6	13
31-40	60-89	-	2	-	1	3
	90-99	1	-	-	-	1
	100-109	-	1	-	1	2
	110-129	2	1	-	1	4
	Total	3	4	0	3	10
41-65	60-89	-	-	1	-	1
	90-99	1	1	-	-	2
	100-109	2	1	-	-	3
	110-129	1	-	3	-	4
	Total	4	2	4	0	10
Total		15	9	4	12	40

Table III-3a

Pilot Group - Outcome Status

Age	Sex	Now Working	Not Now Working		No Data On Follow-up	Total
			Has Worked	Never Worked		
	Male	6	1	1	3	11
	Female	3	4	4	1	12
14-19						
	Total	9	5	5	4	23
	Male	2	3	3	4	12
	Female	4	7	1	1	13
20-23						
	Total	6	10	4	5	25
	Male	14	2	-	2	18
	Female	2	3	-	1	6
24-33						
	Total	16	5	0	3	24
	Male	7	4	2	3	16
	Female	4	2	3	4	13
34-63						
	Total	11	6	5	7	29
Total		42	26	14	19	101

Table III-3b

Cross-Validation Group - Outcome Status

Age	Sex	Now Working	Not Now Working		No Data On Follow-up	Total
			Has Worked	Never Worked		
	Male	-	-	-	2	2
	Female	3	1	-	1	5
14-20						
	Total	3	1	0	3	7
	Male	2	-	-	3	5
	Female	3	2	-	3	8
21-30						
	Total	5	2	0	6	13
	Male	3	3	-	1	7
	Female	-	1	-	2	3
31-40						
	Total	3	4	0	3	10
	Male	4	-	4	-	8
	Female	-	2	-	-	2
41-65						
	Total	4	2	4	0	10
Total		15	9	4	12	40

Table III-4a

Pilot Group - Outcome Status

Marital Status	Sex	Now Working	Not Now Working		No Data On Follow-up	Total
			Has Worked	Never Worked		
Single	Male	14	8	5	9	36
	Female	10	15	7	4	36
	Total	24	23	12	13	72
Married	Male	11	2	1	2	16
	Female	1	-	-	-	1
	Total	12	2	1	2	17
Separated	Male	1	-	-	-	1
	Female	2	-	1	-	3
	Total	3	0	1	0	4
Divorced	Male	3	-	-	1	4
	Female	-	1	-	2	3
	Total	3	1	0	3	7
Widowed	Male	-	-	-	-	0
	Female	-	-	-	1	1
	Total	0	0	0	1	1
Total		42	26	14	19	101

Table III-4b

Cross-Validation Group - Outcome Status

Marital Status	Sex	Now Working	Not Now Working		No Data On Follow-Up	Total
			Has Worked	Never Worked		
Single	Male	4	-	1	5	10
	Female	6	4	-	6	16
	Total	10	4	1	11	26
Married	Male	2	2	3	1	8
	Female	-	-	-	-	0
	Total	2	2	3	1	8
Separated	Male	1	1	-	-	2
	Female	-	-	-	-	0
	Total	1	1	0	0	2
Divorced	Male	2	-	-	-	2
	Female	-	2	-	-	2
	Total	2	2	0	0	4
Widowed	Male	-	-	-	-	0
	Female	-	-	-	-	0
	Total	0	0	0	0	0
Total		15	9	4	12	40

Table III-5

Some Relationships of Interest Between Client Characteristics and Outcome

a) Months out of MRC and Outcome - Pilot Group Only

Months Out of MRC	Outcome		
	Working at Follow-up	Not Working at Follow-up	Total
16-28	15	6	21
29-38	27	34	61
Total	42	40	82

b) Marital Status and Outcome - Pilot and Cross-Validation Groups

Marital Status	Outcome		
	Working at Follow-up	Not Working at Follow-up	Total
Single	34	40	74
Married ????	23	13	36
Total	57	53	110

c) Sex and Outcome - Pilot and Cross-Validation Groups

Sex	Outcome		
	Working at Follow-up	Not Working at Follow-up	Total
Male	38	23	61
Female	19	30	49
Total	57	53	110

Table III-5 (Continued)

Some Relationships of Interest Between Client Characteristics and Outcome

d) I.Q. and Outcome - Pilot and Cross-Validation Group Males Only

I.Q.	Outcome: Working at Follow-up	Outcome: Not Working at Follow-up	Total
"Average" (90-109)	25	15	40
"Extreme" (60-89 & 110-129)	13	26	39
Total	38	41	79

e) Years of Education and Outcome - Pilot and Cross-Validation Groups

Years of Education	Outcome: Working at Follow-up	Outcome: Not Working at Follow-up	Total
5-11	26	18	44
12-16	31	35	66
Total	57	53	110

APPENDIX IV

MEDICAL CODING FOR REHABILITATION CLIENTS

Finding useful methods of characterization of the medical status of clients in vocational rehabilitation centers has always been difficult--and no well-accepted diagnostic coding scheme for such problems exist. Yet, the client's status, vis-a-vis such problems can make the difference between successful and unsuccessful rehabilitation. The project staff gave particular attention to the development of a short, easy medical coding scheme which would have some predictive validity for the kinds of populations seen in comprehensive rehabilitation centers. As will become evident later, the attempt had a fair amount of success and should be useful in future research in this area.

The coding scheme developed was essentially a dichotomous one, coding a yes ("1"), no ("0", "no information" or "no decision possible" codings in answer to the following questions:

1. Is any medical disability of the client progressive? (versus, of course, whether it is stable).

2. Is any disability congenital?

3. Is there a motor disability?

4. Is there mental retardation?

5. Is there an orthopedic disability?

6. Is there psychiatric involvement?

7. Is there a sensory involvement?

The rules for this coding scheme were developed by one of the authors, and were the following:

1. Progressive versus Stable. Here, if the disease was known a priori (e.g., Parkinson's Disease, Multiple Sclerosis) to be progressive, it was so coded; otherwise, some reference to the progressive nature of a diagnosed disability had to be given. Or, to be recorded as stable, the same kind of indications had to be present. In general, such disabilities as residuals of poliomyelitis, mental retardation, cerebral palsy, and traumatic hemiplegias were so categorized. In this coding, because of lack of decision information, or inability to make the decision on medical grounds, such impairments as the following could not be classified: chronic ulcers, seizures, psychiatric disorders, obesity, emphysema, chronic brain syndrome due to alcoholism, back injuries, CNS disease, arthritis, glomerulonephritis.

2. Congenital versus non-Congenital. Any impairment which was by definition congenital, or was present or noticed at birth, was classified as congenital. This included some convulsive disorders, congenital orthopedic deficiencies, dwarfism and the like. Where symptoms of a disorder had appeared at a definite time, it was coded as non-congenital. When etiology is unknown, or there is not enough information, the coding was "indeterminate" - which included, for this project, some mental retardations, MS,

Friedrich's staxia, and CNS disease.

3. Motor versus non-Motor. Problems in movement, muscular or neuro-muscular weakness or poor control (e.g., back strain, spasticity, paralysis, and arthritis) were included as motor.

4. Retardation versus non-Retardation. Any client whose highest recorded IQ was less than 80 was listed as retarded.

5. Orthopedic versus non-Orthopedic. Any involvement of bone structure, bone growth, fractures, deformities, amputations, arthritis, dwarfism, back problems, and like, were included.

6. Psychiatric versus non-Psychiatric. Any history of mental illness or hospitalization for mental illness was included.

7. Sensory versus non-Sensory. Any sensory disability (e.g., blindness, tactile disorders) was included.

It should be noted that a number of multiple codings for one particular diagnosis could occur. For example, back problems and arthritis would be categorized as motor and orthopedic.

As it turns out, for both experimental and control groups, the progressive-stable and sensory-non-sensory codings included only a few cases (8 cases in the experimental group and none in the control group were listed as progressive, while ten cases in the experimental group and only one in the control group were coded as sensory). If we then consider only the four other disability codings (psychiatric, retardation, motor, and orthopedic - lumping the "indeterminate" and "non" codings), and attach the congenital description, when needed, only one-seventh of the experimentals (21 of 141) were without some disability coding (see Table III-I). A little less than half of the group (67 cases) was coded as single-disability. The remainder, about three-eighths of the experimental group (53 cases) were coded multiple-disability with a slightly heavier weighting toward psychiatric cases and toward single-disability cases the only differences of note.

Table IV-1

Distribution of Disability-Combinations Among Experimental and Control Groups

	Experimental	Control
No Disability Listed	21	4
Psychiatric Only	36	11
Orthopedic Only	3	
Congenital Orthopedic Only	2	
Orthopedic and Psychiatric	1	
Retardation Only	9	
Congenital Retardation Only	3	
Retardation and Psychiatric	1	
Congenital Retardation and Psychiatric	1	
Motor Only	26	6
Congenital Motor Only	9	2
Psychiatric and Motor	5	
Congenital Motor and Psychiatric	1	
Motor and Orthopedic	14	4
Congenital Motor and Orthopedic	4	1
Psychiatric, Motor, and Orthopedic	1	
Motor and Retardation	1	
Congenital Motor and Retardation	2	
Motor Retardation and Psychiatric	1	
Motor Retardation and Orthopedic		1
Total	141	29

Table IV-2

Medical Coding Data, Bivariate Distribution of Congenital Nature of Handicap and All Other Handicaps, Pilot and Cross Validation Groups

Coded As:		Pilot Group			Cross-Validation Group		
		Coded as Congenital?			Coded As Congenital?		
		Yes	No	Total	Yes	No	Total
Progressive in Nature?	Yes	1	6	7	-	1	1
	No	21	73	94	5	34	39
	Total	22	79	101	5	35	40
Motor?	Yes	15	38	53	1	10	11
	No	7	41	48	4	25	29
	Total	22	79	101	5	35	40
Mental Retardation?	Yes	5	8	13	1	4	5
	No	17	-	88	4	31	35
	Total	22	79	101	5	35	40
Orthopedic?	Yes	5	15	20	1	4	5
	No	17	64	81	4	31	35
	Total	22	79	101	5	35	40
Psychiatric?	Yes	2	26	28	-	19	19
	No	20	53	73	5	16	21
	Total	22	79	101	5	35	40
Sensory?	Yes	1	6	7	1	2	3
	No	21	73	94	4	33	37
	Total	22	79	101	5	35	40

CONFIDENTIAL

MINNEAPOLIS REHABILITATION CENTER
1900 Chicago Avenue 333-2335

CONFIDENTIAL

FOLLOW-UP QUESTIONNAIRE

Please complete this form. Most of the questions have several choices with a short line after each choice. Wherever you see this, please make an X on the line after the answer which is the most correct for you. Mark only one answer. There are a few questions which ask for an explanation in your own words but most of these questions can be answered simply by marking X.

1. Are you now working? ____Yes ____No If no, why not?_____

_____

2. Do you work? ____Full-time ____Part-time ____It varies

3. What is your pay? $_____
   ____Is this per hour? ____Is it take-home pay?
   ____Is this per week? ____Is it before taxes pay?
   ____Is this per month?

4. How many jobs have you had since leaving MRC?
   ____None ____One ____Two to Four ____Five or more

5. Please add up the amount of time you worked on all the jobs you held since you left MRC and write that total time here.
   _____months

6. Have you moved since leaving MRC? ____No ____Yes - How many times?
   ____Once ____4 times
   ____Twice ____5 times
   ____3 times ____More than 5

7. Has your marital status changed since leaving MRC? ____No ____Yes
   ____Got married ____Got separated
   ____Got divorced ____Wife or husband died

8. Have you been laid up with illness, injury or for any other reason for more than two weeks since you left MRC? ____No ____Yes - For how long?
   ____2 to 4 weeks ____4 to 6 weeks ____6 weeks to 2 months
   ____2 to 3 months ____3 to 4 months ____4 to 6 months
   ____6 to 9 months ____9 to 12 months ____More than 1 year

9. Since you left MRC, have things: ____gotten better
   ____gotten worse
   ____stayed the same

10. What, if anything, did you like about MRC:_____

_____

11. What, if anything, did you dislike about MRC?_____

_____

12. Have you any other comments?_____

Name _____ Number _____

Address _____

City & State _____

The form on the following pages is to be used as part of a study on people who came to the Minneapolis Rehabilitation Center. Only the people doing the study will see this form, and your name will not appear on it.

Please write your correct name and address at the top of this page. When we receive your completed form you will be given a number in order to identify your form to the people doing the study and this page will be torn off and kept in a locked file apart from the form itself. Your number will be known only to these people doing the study and will therefore keep unauthorized people from knowing who filled out your form.

Please answer all of the questions truthfully. None of the statements you make will affect your position with your family, employer or anywhere else that you might not want people to know what you said.

Thank you for your cooperation.

CONFIDENTIAL

MINNEAPOLIS REHABILITATION CENTER
1900 Chicago Avenue 333-2335

CONFIDENTIAL

FOLLOW-UP QUESTIONNAIRE

_____
Number

Please complete this form. Most of the questions have several choices with a short line after each choice. Wherever you see this, please make an X on the line after the answer which is the most correct for you. Mark only one answer. There are a few questions which ask for an explanation in your own words but most of these questions can be answered simply by marking X.

1. Are you now working: ___Yes ___No If no, why not? _____

_____

2. If the reason is you cannot get a job, please explain why you cannot get one.

_____

_____

3. If you are working, where? (MRC will NOT contact any employers named) _____

_____

4. What kind of business, company or organization is it? _____

_____

5. What is your job title? _____

6. What exactly do you do on your job? _____

_____

7. Do you work? ___full-time ___part-time ___it varies

8. If you work part-time or if your time varies, do you usually work:
___more than 40 hours per week
___about 40 hours per week
___30 to 40 hours per week
___20 to 30 hours per week
___10 to 20 hours per week
___less than 10 hours per week

9. What is your pay? $_____
___Is this per hour?
___Is this per week?
___Is this per month?
___Is it take-home pay?
___Is it before taxes pay?

10. If you are working, how much is your present job like the kind of work that you planned for during counseling?
___Exactly what I planned for in counseling
___Much like what I planned for in counseling
___Somewhat the same
___Not at all the same

11. If you received training, did you use your training on any past or present job?
___Very much ___Somewhat ___Very little ___No training received

12. What do you like most about your job and what do you like least? (Remember MRC will NOT contact any employers.)

Like Most	Like Least
___The work	___Tne work
___Hours	___Hours
___Pay	___Pay
___The boss	___The boss
___Co-workers	___Co-workers
___Easy to get there	___Hard to get there
___Other (explain) _____	___Other (explain _____
_____	_____
_____	_____

13. Are you satisfied with your job? ___Yes ___No

14. Are you planning to change jobs in the near future? ___Yes ___No

15. How many jobs have you had since living MRC?
___None ___1 ___2 to 4 ___5 or more

16. Since you left MRC what is the longest and shortest time you stayed on your jobs?
Longest: ___months Shortest: ___months

17. Please add up the amount of time you worked on all the jobs you held since you left MRC and write that total time here.
_____ months

18. Have you moved since leaving MRC? ___No ___Yes How many times?

___Once	___4 times
___Twice	___5 times
___3 times	___More than 5

19. Has your marital status changed since leaving MRC? ___No ___Yes How?

___Got married	___Got separated
___Got divorced	___Wife or husband died

20. Have you been laid up with illness, injury or for any other reason for more than two weeks since you left MRC? ___No ___Yes For how long?

___2 to 4 weeks	___4 to 6 weeks	___6 weeks to 2 months
___2 to 3 months	___3 to 4 months	___4 to 6 months
___6 to 9 months	___9 to 12 months	___More than 1 year

21. What was the reason? _____

_____

22. Since you left MRC, have you: ___bought a car ___sold your car
___lost your car ___no change

23. Since you left MRC, have you: ___bought a house ___sold your house
___lost your house ___no change

24. Since you left MRC, have things: ___gotten better
___gotten worse
___stayed the same

25. If you checked "gotten better" or "gotten worse", please explain how.

_____

26. Who referred you to MRC? ___(1) Division of Vocational Rehabilitation (DVR)
___ (2) State Services for the Blind ___(3) Employment Service
___ (4) County Welfare ___(5) Don't know

___ (6) Other (Explain) _____

27. Why did you originally come to MRC? _____

_____

Answer the following statements by marking an X on the line under True if the statement is true for you or by marking an X on the line under False if the statement is false for you. Mark only one line for each item.

			True	False
28.	MRC helped me with:	Getting a job	___	___
29.		Getting a better job	___	___
30.		Keeping a job	___	___
31.		Training	___	___
32.		Learning more about myself	___	___
33.		Feeling better about myself	___	___
34.		My family	___	___
35.		Friends	___	___
36.		Agencies (like DVR, County, etc.)	___	___
37.		Other	___	___
		(If True for "other", please explain _____		
		_____		
38.	MRC did not help me enough		___	___
		(If True, what should MRC have done to help you more?)		
		_____		
39.	Another agency helped me more than MRC did		___	___
		(If True, which agency?) _____		
		_____		
40.	I think MRC can help most of the people that go there		___	___
While I was at MRC, it was sometimes hard to get along with:				
41.	1 or 2 of the MRC staff members.		___	___
42.	Almost all of the MRC staff members		___	___

		True	False
43.	1 or 2 of the other clients	___	___
44.	Almost all of the other clients	___	___
45.	1 or 2 members of my family	___	___
46.	Almost my whole family	___	___

47. What, if anything, did you like about MRC? _____

_____

48. What, if anything, did you dislike about MRC? _____

_____

49. Have you any other comments? _____

_____

_____

Thank you for your cooperation in completing this questionnaire.